

『現代日本語書き言葉均衡コーパス』におけるサンプル構成比の算出法：現代日本語書き言葉の文字数調査

著者	丸山 岳彦, 秋元 祐哉
ページ	1-75
発行年	2007-03-31
シリーズ	国立国語研究所内部報告書 ; LR-CCG-06-02
URL	http://doi.org/10.15084/00002838

『現代日本語書き言葉均衡コーパス』における サンプル構成比の算出法

—現代日本語書き言葉の文字数調査—

丸山 岳彦・秋元 祐哉

平成19年3月

大規模汎用日本語データベースの構築とその活用に関する調査研究

©2007 独立行政法人国立国語研究所

国立国語研究所 内部報告書 (LR-CCG-06-02)

『現代日本語書き言葉均衡コーパス』 におけるサンプル構成比の算出法

—現代日本語書き言葉の文字数調査—

丸山 岳彦 秋元 祐哉

平成19年 3月

大規模日本語汎用データベースの構築とその活用に
関する調査研究

©2007 独立行政法人国立国語研究所

目次

第 I 部	本編	3
第 1 章	「現代日本語書き言葉の文字数調査」の目的	5
1.1	BCCWJ の全体構成	5
1.1.1	BCCWJ を構成する 3 つのサブコーパス	5
1.1.2	BCCWJ で提供される 2 種類のサンプル	6
1.2	生産実態サブコーパスの設計方針	7
1.3	「現代日本語書き言葉の文字数調査」の立案	8
第 2 章	「現代日本語書き言葉の文字数調査」の方法	9
2.1	母集団の定義と層別	9
2.1.1	書籍の母集団の定義と層別	9
2.1.2	雑誌の母集団の定義と層別	11
2.1.3	新聞の母集団の定義と層別	12
2.2	調査方法	14
2.2.1	計測ページの抽出	14
2.2.2	計測の対象とする文字要素	16
2.2.3	計測の方法	17
第 3 章	「現代日本語書き言葉の文字数調査」の結果	19
3.1	調査結果	19
3.1.1	書籍の文字数調査の結果	19
3.1.2	雑誌の文字数調査の結果	20
3.1.3	文字数調査の結果 (3) — 新聞	23
3.2	調査のまとめ	25
3.3	サンプル構成比の算出	30
3.3.1	サンプル構成比の求め方	30
3.3.2	必要サンプル数の算出	30
第 II 部	資料編	33
第 4 章	「現代日本語書き言葉の文字数調査」資料集	35

目 次

2.1	「適切性条件」による書籍の絞り込みの例	10
2.2	生産実態サブコーパス「書籍」の母集団（NDC による層別）	11
2.3	生産実態サブコーパス「雑誌」の母集団（分野による層別）	12
2.4	生産実態サブコーパス「新聞」の母集団（紙種による層別）	13
2.5	書籍（NDC「0. 総記」）の判型ごとの冊数	15
2.6	書籍の計測対象として抽出された冊数	15
2.7	雑誌（分野「1. 総合」）の判型ごとの冊数	15
2.8	雑誌の計測対象として抽出された冊数	16
2.9	目視計測と OCR 計測の結果の誤差（一部）	17
2.10	目視計測を実施した冊数と OCR 計測を実施した冊数	18
3.1	書籍の文字数計測結果（一部）	19
3.2	書籍の総文字数推計基準表	21
3.3	書籍の文字数調査のまとめ — NDC 別の推計総文字数	21
3.4	雑誌の文字数推計結果（一部）	22
3.5	雑誌の総文字数推計基準表	22
3.6	雑誌の文字数調査のまとめ — 分野別の推計総文字数	23
3.7	朝日新聞朝刊の面種ごとの面積と計測した文字数（一部）	24
3.8	朝日新聞 1 週間分の面種ごとの面積と推計総文字数（一部）	24
3.9	朝日新聞朝刊 1 週間分の文字数・ページ数と、キャラクタ密度の算出	25
3.10	新聞の文字数調査のまとめ — 新聞タイトル別の推計総文字数	26
3.11	2001 年から 2005 年の書籍・雑誌・新聞の推計総文字数	26
3.12	2001 年から 2005 年の書籍・雑誌・新聞の推計総文字数 (2)	30
3.13	生産実態サブコーパス全体のサンプル構成比	31
4.1	生産実態サブコーパス「書籍」の母集団（発行年・NDC による層別）	35
4.2	2001 年から 2005 年の各年における雑誌協会加盟社のリスト	36
4.3	生産実態サブコーパス「雑誌」の母集団（発行年・分野による層別）	37
4.4	生産実態サブコーパス「新聞」の母集団（紙種・新聞タイトルによる層別）	38
4.5	書籍の NDC・判型ごとの冊数と累計	39
4.6	雑誌の分野・判型ごとの冊数と累計	40
4.7	目視計測と OCR 計測の結果の誤差 — 書籍	41
4.8	目視計測と OCR 計測の結果の誤差 — 雑誌	42

4.9 書籍の文字数計測結果 (1/5)	43
4.10 書籍の文字数計測結果 (2/5)	44
4.11 書籍の文字数計測結果 (3/5)	45
4.12 書籍の文字数計測結果 (4/5)	46
4.13 書籍の文字数計測結果 (5/5)	47
4.14 書籍の文字数調査のまとめ — 発行年・NDC・判型別の冊数	48
4.15 書籍の文字数調査のまとめ — 発行年・NDC・判型別のページ数	50
4.16 書籍の文字数調査のまとめ — 発行年・NDC・判型別の推計総文字数	52
4.17 雑誌の文字数計測結果	54
4.18 雑誌の文字数調査のまとめ — 発行年・分野・判型別の冊数	56
4.19 雑誌の文字数調査のまとめ — 発行年・分野・判型別のページ数	58
4.20 雑誌の文字数調査のまとめ — 発行年・分野・判型別の推計総文字数	60
4.21 新聞の文字数計測結果 (朝日新聞朝刊)	64
4.22 新聞の文字数計測結果 (毎日新聞朝刊)	65
4.23 新聞の文字数計測結果 (読売新聞朝刊)	66
4.24 新聞の文字数計測結果 (日本経済新聞朝刊)	67
4.25 文字数計測を行った新聞の日付と、面積を計測した新聞の日付	68
4.26 全国紙 4 紙の面種のまとめあげ	69
4.27 新聞 1 週間分の推計総文字数 (朝日新聞)	70
4.28 新聞 1 週間分の推計総文字数 (毎日新聞)	71
4.29 新聞 1 週間分の推計総文字数 (読売新聞)	72
4.30 新聞 1 週間分の推計総文字数 (日本経済新聞)	73
4.31 全国紙 4 紙 1 週間分の文字数・ページ数と、キャラクタ密度の算出	74
4.32 新聞の文字数調査のまとめ — 新聞タイトル (朝夕刊) 別の推計総文字数	75

目 次

1.1	BCCWJ の構成	5
2.1	5 年間における書籍の冊数・ページ数 (1)	10
2.2	5 年間における書籍の冊数・ページ数 (2)	10
2.3	5 年間における雑誌のタイトル数・冊数・ページ数	12
2.4	5 年間における新聞のタイトル数・冊数・ページ数	13
2.5	新聞の母集団となるタイトル一覧	13
2.6	2003 年に発行された書籍・雑誌・新聞の総数	14
3.1	書籍の判型別キャラクタ密度と回帰直線	20
3.2	書籍の総文字数構成比	27
3.3	雑誌の総文字数構成比	28
3.4	新聞の総文字数構成比	29
3.5	固定長サンプル (1,000 万語分) の構成比	31
4.1	目視計測で用いた新聞のコピーの例 (その 1)	62
4.2	目視計測で用いた新聞のコピーの例 (その 2)	63

はじめに

本報告書は、『現代日本語書き言葉均衡コーパス』の設計（コーパスデザイン）の一環として実施した「現代日本語書き言葉の文字数調査」について示すものである。『現代日本語書き言葉均衡コーパス』は、1976年から2005年の30年間に生産された現代日本語の書き言葉を収録する、1億語規模のコーパスである。独立行政法人国立国語研究所研究開発部門言語資源グループを中心に、5カ年計画で2006年4月から構築を開始した。

『現代日本語書き言葉均衡コーパス』の構築にあたり、そのコーパスデザインの一環として、我々は、いくつかの基礎調査を行った。本報告書で述べる「現代日本語書き言葉の文字数調査」も、その一つである。この調査では、2001年から2005年までに国内で発行された書籍・雑誌・新聞を対象として、そこにどれだけの文字が含まれているかを推計する。調査によって推計された総文字数の分布は、本編中で解説する「生産実態サブコーパス」と呼ばれるサブコーパスに格納する書籍・雑誌・新聞のサンプル構成比として採用する。すなわち、総文字数の分布を書き言葉の「生産力」を反映するものと見なすことにより、書き言葉の生産実態をできる限り忠実に反映するコーパスの構築を目指そうとするのである。

「現代日本語書き言葉の文字数調査」は、2005年4月から2006年4月にかけて、国立国語研究所研究開発部門言語資源グループに所属する以下の者が中心となり実施した。

前川喜久雄，山崎誠，柏野和佳子，丸山岳彦，秋元祐哉，稲益佐知子，吉田谷幸宏

調査の実施にあたっては、市原乃奈，井上陽子，大矢内夢子，立花幸子，田中美恵子，田中弥生，長門美帆子，西野文，服部紀子，保田祥，吉田奈央，李承赫らが、これを助けた。調査結果の集計は、主として秋元祐哉が担当した。本報告書の執筆は、丸山岳彦と秋元祐哉の2名が担当した。

第I部
本編

第1章 「現代日本語書き言葉の文字数調査」 の目的

本章では「現代日本語書き言葉の文字数調査」の目的について示す。まず『現代日本語書き言葉均衡コーパス (Balanced Corpus of Contemporary Written Japanese; 以下 BCCWJ と記す)』の構成を確認した後、「現代日本語書き言葉の文字数調査」を実施した経緯と、調査の目的について述べる。

1.1 BCCWJ の全体構成

はじめに、BCCWJ の構成について確認しておく¹。BCCWJ は、過去 30 年の間に生産された現代日本語の書き言葉を収録するコーパスである。書籍・雑誌・新聞・白書などを主な対象として書き言葉のサンプルを抽出し、全体で 1 億語規模のコーパスとなることを目指す。

1.1.1 BCCWJ を構成する 3 つのサブコーパス

BCCWJ 全体は、「生産実態サブコーパス」「流通実態サブコーパス」「非母集団サブコーパス」と呼ぶ、性格の異なる 3 種類のサブコーパスから構成される。このうち生産実態サブコーパスと流通実態サブコーパスでは、母集団を厳密に定義し、層別ランダムサンプリングによってサンプルを抽出する。これを実現するために必要となるのは、書き言葉の生産実態および流通実態に関する基礎調査と、母集団の定義、サンプリング台帳の作成、そして一定の手続きに基づくサンプリングの実施である。

BCCWJ の構成全体を図示すると、図 1.1 のようになる。各サブコーパスの概要を、以下に示す。

生産実態サブコーパス 書籍、雑誌、新聞 2001～2005年 約3500万語 固定長+可変長	流通実態サブコーパス 書籍 1976～2005年 約3000万語 固定長+可変長
非母集団サブコーパス 白書、法律、国会会議録、ベストセラー、 教科書、WWW上のテキスト ... 1976～2005年 2001～2005年 約3500万語 可変長(一部 固定長+可変長)	

図 1.1: BCCWJ の構成

¹ BCCWJ の設計全般については、山崎 (2007) を参照のこと。

生産実態サブコーパス 生産実態サブコーパスは、書き言葉の生産力という側面に着目するサブコーパスである。2001年から2005年の間に国内で出版されたすべての書籍・雑誌・新聞を母集団として、ランダムサンプリングによって得られる約3,500万語分のデータを収める。書き言葉が実際に出版された結果をできる限り忠実に反映することで、5年間における書き言葉の生産に関するありさまを捉えるサブコーパスを作ることとする。

流通実態サブコーパス 流通実態サブコーパスは、書き言葉の流通・流布という側面に着目するサブコーパスである。2005年の時点において東京都内の公共図書館に所蔵されている書籍（ただし1976年から2005年の30年間に発行されたもの）を母集団として、ランダムサンプリングによって得られる約3,000万語分のデータを収める。書き言葉（書籍）が流通している度合いを公共図書館の所蔵状況によって近似的に把握し、世の中に広く流布・遍在している書き言葉のありさまを捉えることを目的とする。

非母集団サブコーパス 非母集団サブコーパスは、生産・流通という側面からは捉えきれない書き言葉のうち、研究遂行上必要と思われる種類の書き言葉を収めるサブコーパスである。白書、教科書、法律、国会会議録、WWW上の文書などを対象として、約3,500万語分のデータを収める。

1.1.2 BCCWJで提供される2種類のサンプル

次に、BCCWJで提供される2種類のサンプルについて述べる。我々は、BCCWJに収録するサンプルが備えるべき条件として、以下の方針を立てた。

- 統計的に厳密な言語調査に耐え得るよう、母集団からの抽出比を重視した設計にする。
- 文体研究・テキスト研究に耐え得るよう、ある程度の文脈を確保した設計にする。

これらの方針に対処するため、「固定長サンプル」「可変長サンプル」という、異なる長さを持つ2種類のサンプルを設計した。両者は、母集団からの抽出方法という点において異なる仕様を持つ。

固定長サンプル 「固定長サンプル」は、母集団に含まれるすべての文字に対して等確率を与えた上で、ある1文字をランダムに指定し、その文字を始点として1,000文字の範囲を抽出するサンプルである²。すべての文字に対して等確率を与えるために、母集団に含まれる文字の総数をあらかじめ推計しておく必要がある。母集団（＝推計された総文字数）からの抽出比が明確である点で、基本語彙表や漢字表の作成、語彙・文字調査など、統計的な言語研究に向く。また、母集団の層的・量的な構造が忠実に反映される点で、バランストコーパスを構成するサンプルとしての性格を強く持つ。

可変長サンプル 「可変長サンプル」は、固定長サンプルと同様、母集団に含まれるすべての文字に対して等確率を与えた上で、ランダムに指定した1文字を含む言語的な構造のまとまり（「章」や「節」など、ただし1万字を上限とする）を抽出するサンプルである。文章・談話としてのまとまりを重視したサンプルであるため、テキストの論理構造の把握や文脈の分析、文体の調査などに向く。

² 実際に1,000文字としてカウントするのは、漢字、仮名、数字、アルファベットのみである。句読点や記号類は数えない。

可変長サンプルは、BCCWJ を構成する 3 つのサブコーパスのすべてに対して提供される。一方の固定長サンプルは、統計的な調査を行う可能性の高い部分、すなわち、生産実態サブコーパス、流通実態サブコーパス、および、非母集団サブコーパスの一部（白書など）に対して、可変長サンプルと同時に提供される。

1.2 生産実態サブコーパスの設計方針

本報告書で取り上げるのは、3 つのサブコーパスのうち、生産実態サブコーパスである。以下では、生産実態サブコーパスのコーパスデザイン、およびサンプリングの方針について述べる。

生産実態サブコーパスを設計するにあたり、我々はまず、書き言葉の語彙調査を実施することを主たる目的として、2001 年から 2005 年の間に国内で発行されたすべての書籍・雑誌・新聞から合計 1,000 万語分の固定長サンプルを抽出することを計画した。1,000 万語という数値は、統計的な語彙調査のために十分なサンプルサイズを見積もった結果である。

1,000 万語の内訳については、5 年間に発行された書籍・雑誌・新聞に含まれる総文字数を推計し、その比を採用することにした。例えば、書籍・雑誌・新聞の総文字数の推計値が「5:3:2」の割合であれば、書籍から 500 万語、雑誌から 300 万語、新聞から 200 万語を取得する、という具合である。これは、書き言葉の生産力を文字の数によって近似的に捉えることにより、サブコーパス全体が書き言葉の生産実態のいわば縮図となることを意図したものである。

また、従来行われてきた語彙調査の結果などを参考にして、1 語あたりの平均長を約 1.7 文字と試算した。ここから、1,000 文字を 1 サンプルとする固定長サンプルで 1,000 万語分を達成するためには、17,000 サンプルが必要となるという見積もりを立てた。

さらに、固定長サンプルを取得するために指定した 1 文字を利用して、可変長サンプルを同時に取得することにした。複数の資料に基づいて可変長サンプルの平均長を試算したところ、書籍で平均 3,900 文字、雑誌で平均 3,000 文字、新聞で平均 1,000 文字という結果が得られた。可変長サンプル全体を合計するとどれだけのサイズになるかは、固定長サンプル 1,000 万語の構成比が算出されてから決まることになる。

以上のようなコーパスデザインに基づき、生産実態サブコーパスでのサンプリングを、概略、次の手順で進めることにした。

まず、2001 年から 2005 年までに国内で発行された書籍・雑誌・新聞を調査し、客観的に把握できる形で母集団を厳密に定義する。次に、書籍・雑誌・新聞それぞれを「層」と見なし、さらにそれらを下位の層に分割した上で、各層に含まれる文字の数を推計する。これにより、生産実態サブコーパスの母集団を文字数という量的な側面から把握した上で、総文字数の比に応じて各層の構成比を決定することができる。さらに、層別ランダムサンプリングを行うためのサンプリング台帳を作成する。これは、各層に含まれるすべての文字に対して等確率を与え、その中の 1 文字をランダムに指定するように設計する。指定された 1 文字は「サンプル抽出基準点」となり、この点を基準として、固定長サンプルでは 1,000 文字の範囲が、可変長サンプルではその文字を含む「章」「記事」などの範囲が、それぞれサンプルとして抽出されることになる。あとは、サンプリング台帳で指定されている実際の書籍などを手に取り、一定の手順に従って、印刷紙面からサンプルを抽出する作業を個別に進めていけばよい。

1.3 「現代日本語書き言葉の文字数調査」の立案

以上のような手順によってサンプリングを実施するためには、前もって、次の4つの点を明らかにしておく必要がある。

1. 母集団の定義 — 母集団をどのように定義するか。それをどのように層別するか。
2. サンプル構成比の決定 — 各層に含まれる総文字数をどのように推計し、固定長サンプル1,000万語分の構成比を算出するか。
3. サンプリング台帳の作成 — どのような手順により1文字をランダムに指定するか。
4. サンプリング手法の開発 — 指定された1文字を基準として、固定長サンプル・可変長サンプルをどのように取得するか。

このうち「1. 母集団の定義」は、次章で詳しく述べるように、書籍については国立国会図書館の蔵書目録を、雑誌・新聞についてはそれぞれ『雑誌新聞総かたろぐ』『全国新聞ガイド』を用いて、定義を行った。そして、「2. サンプル構成比の決定」を明らかにするために企画・立案されたのが、本報告書で詳しく述べる「現代日本語書き言葉の文字数調査」である。この調査では、1. で定義した書籍・雑誌・新聞の母集団に含まれる文字数を推計することにより、生産実態サブコーパスの固定長サンプル1,000万語分のサンプル構成比を算出することを目的とする。サンプル構成比が算出された後は、「3. サンプリング台帳の作成」「4. サンプリング手法の開発」について検討する必要があるが、これらについては本稿では触れず、別稿に委ねる³。

そもそも、2001年から2005年の間にどれだけの書き言葉が生産されたのか、そこにはどれだけの数の文字が含まれているのか、といった調査は、一見、コーパスの構築とは無関係であるように見える。しかしながら、文字数という側面から書き言葉の生産実態を忠実に反映するようにコーパスを構築しようとする限り、あるいは、コーパスに含まれるサンプルの、母集団からの抽出比を明確にしようとする限り、母集団に含まれる総文字数の推計は、欠くことのできない手順であると言える。

³ サンプリング台帳の作成、およびサンプリング手法の概要については、丸山・柏野・稲益他(2007)、丸山・柏野・山崎他(2007)を参照のこと。

第2章 「現代日本語書き言葉の文字数調査」の方法

本章では「現代日本語書き言葉の文字数調査」の方法および手順について示す。まず、BCCWJ 全体の設計に関して、生産実態サブコーパスが対象とする書籍・雑誌・新聞の母集団の定義について述べ、この母集団をさらに下位の層に層別する方法について述べる。次に、「現代日本語書き言葉の文字数調査」として、得られた母集団の定義と層別の結果を用いて、各層からランダムに抽出したページに含まれる文字数を計測し、それらを集計することで、母集団全体の文字数を推計する。

以下では、母集団の定義と層別の方法について述べた後、「現代日本語書き言葉の文字数調査」を実施するための準備として、調査対象および調査方法をどのように定めたかを示す。

2.1 母集団の定義と層別

はじめに、2001 年から 2005 年までの書き言葉の生産実態を調査し、生産実態サブコーパスが対象とする母集団を定義した結果について、書籍・雑誌・新聞の順に述べる。

2.1.1 書籍の母集団の定義と層別

まず、生産実態サブコーパス「書籍」の母集団の定義と、層別の方法について述べる。

2001 年から 2005 年の間に国内で発行された書籍の実態を把握するために、我々は、国立国会図書館の蔵書目録を利用することにした。国内で出版される出版物は、国立国会図書館法により、すべて国立国会図書館に納本されることになっている。そこで、この蔵書目録を使えば、国内における書籍の生産実態をほぼ網羅的に捉えることができると考えた。

我々は、国立国会図書館・日本図書館協会からの協力を得て、国立国会図書館の蔵書目録を電子化した市販のデータ「J-BISC (Japan Biblio disc)」をリスト化した。このリストは、国内で発行された書籍の書誌情報(タイトル、著者、発行者、発行年、ページ数、NDC など)から構成される。リストを集計した結果、2001 年から 2005 年までに発行された書籍の冊数・ページ数の合計は、図 2.1 のようになった¹。

ただしこの中には、漫画、写真集、人名録のように言語表現が主体でないものや、楽譜、地図、マイクロフィルム、電子出版物(CD-ROM)など、書き言葉コーパスを作るという我々の意図からすると不適切な内容をもつものも多く含まれている。さらに、極端にページ数の少ないもの、非売品など、サンプルに対する著作権処理を施す上で問題となる形態の書籍も多く含まれている。

¹ ただし、561,514 冊中、71,109 冊は、ページ数の欄に「1 冊」「CD-ROM 1 枚」などの情報が入っており、具体的なページ数は得られなかった。

冊数：561,514 冊
ページ数：112,504,115 ページ

図 2.1: 5 年間における書籍の冊数・ページ数 (1)

そこで、コーパスへの収録対象としては不適切な冊をリストから除外するための条件を用意した。これを「適切性条件」と呼ぶ。適切性条件は、リスト中に記載された情報を読み込み、適切でない条件を持つ冊を除外する²。例えば「タイトル」「サブタイトル」「シリーズに関する情報」のいずれかに「写真集」という文字列を含む冊は、コーパスへの収録対象としては不適切な冊として除外される³。適切性条件によって除外される冊数（複数該当）の例を、表 2.1 に示す。

表 2.1: 「適切性条件」による書籍の絞り込みの例

除外条件	除外数
40 ページ以下の書籍	101,417
ページ数の記録がない書籍	67,424
官公庁刊行物のうち非流通物	47,185
学習試験図書	37,780
電子資料、地図資料など	32,635
漫画	24,128
写真集・図画集	10,367
複製、覆刻など	7,832

適切性条件による絞り込みによって、5 年間における書籍の総体は図 2.2 のようになった。この結果を、生産実態サブコーパス「書籍」の母集団として定義した。

冊数：317,117 冊
ページ数：74,911,520 ページ

図 2.2: 5 年間における書籍の冊数・ページ数 (2)

次に、この母集団を層別する基準について示す。ここでは「日本十進分類法 (NDC)」および「発行年」という基準によって層別を行った。「NDC」は日本で用いられている図書館の蔵書分類法であり、書籍の内容により、大きく「0. 総記」「1. 哲学」「2. 歴史」「3. 社会科学」「4. 自然科学」「5. 技術・工学」「6. 産業」「7. 芸術・美術」「8. 言語」「9. 文学」という 10 カテゴリに分類される⁴。我々

² 適切性条件の実体は、1,440 行から成る SQL 文である。

³ ただし、『写真集の作り方』というタイトルの書籍のように、実際は写真集でない冊は、この条件によって除外されないように調整してある。

⁴ 実際の NDC では、「007.637 図形処理ソフトウェア」のように、さらに細かく分類される。

は、J-BISC に付与されている NDC (1 桁目) の 10 分類に加え、NDC が付与されていないレコードを「n (null; 記録なし)」として、合計 11 の層に分類した。これに「発行年」として 2001 年から 2005 年までの 5 分類を重ね合わせ、書籍全体を合計 55 の層に分割した。

生産実態サブコーパス「書籍」の母集団を NDC で層別した際の冊数とページ数について、表 2.2 に示す。

表 2.2: 生産実態サブコーパス「書籍」の母集団 (NDC による層別)

NDC	総冊数		総ページ数	
0. 総記	11,132	(3.51%)	2,859,793	(3.82%)
1. 哲学	18,067	(5.70%)	4,529,329	(6.05%)
2. 歴史	24,624	(7.76%)	6,449,172	(8.61%)
3. 社会科学	62,986	(19.86%)	16,059,116	(21.44%)
4. 自然科学	28,745	(9.06%)	6,771,958	(9.04%)
5. 技術工学	31,377	(9.89%)	6,681,335	(8.92%)
6. 産業	15,332	(4.83%)	3,298,313	(4.40%)
7. 芸術	25,387	(8.01%)	5,153,531	(6.88%)
8. 言語	5,211	(1.64%)	1,196,840	(1.60%)
9. 文学	73,716	(23.25%)	18,888,278	(25.21%)
n. 記録なし	20,540	(6.48%)	3,023,855	(4.04%)
合計	317,117	(100.00%)	74,911,520	(100.00%)

参照 書籍を発行年と NDC で層別した場合の冊数とページ数については、資料編の表 4.1 を参照 (35 ページ)。

2.1.2 雑誌の母集団の定義と層別

次に、生産実態サブコーパス「雑誌」の母集団の定義と、層別の方法について述べる。

「雑誌」という概念を、定期刊行物 (定期的に刊行される冊子) という観点から捉えると、典型的な「雑誌」として想起される月刊誌や週刊誌だけでなく、学会誌や業界誌などの専門誌や、極めて限られた地域だけで流通しているコミュニティ冊子なども含んでしまうため、その全容を把握することが困難になる。そこで、2001 年から 2005 年の間に社団法人日本雑誌協会に加盟していた出版社が発行していた定期刊行物」という条件によって、雑誌を絞り込むことにした。日本雑誌協会の加盟社は国内でも有力な出版社が多く、典型的な「雑誌」の範囲を捉えるのに適切であると判断した。

まず、2001 年から 2005 年の各年における雑誌協会加盟社のリストを作り、対象出版社 102 社を絞り込んだ。さらに、『雑誌新聞総かたろぐ』(メディア・リサーチ・センター発行)を用いて対象出版社が各年に発行した定期刊行物に関する書誌情報を抽出した。その際、新聞、要覧、コミック、非日本語による定期刊行物などは除外した。その結果、2001 年から 2005 年の間に発行された「雑誌」の総体は図 2.3 のようになった。この結果を、生産実態サブコーパス「雑誌」の母集団として定義した。

タイトル数：1,259 タイトル（異なり）

冊数：55,779 冊

ページ数：10,414,955 ページ

図 2.3: 5 年間ににおける雑誌のタイトル数・冊数・ページ数

参照 2001 年から 2005 年の各年における雑誌協会加盟社のリストは、資料編の表 4.2 を参照（36 ページ）。

次に、この母集団を「分野」「発行年」という基準で層別した。「分野」は『雑誌新聞総かたろぐ』で示されている「1. 総合」「2. 教育・学芸」「3. 政治・経済・商業」「4. 産業」「5. 工業」「6. 厚生・医療」という 6 分類である。これに「発行年」として 2001 年から 2005 年までの 5 分類を重ね合わせ、雑誌全体を合計 30 の層に分割した。

生産実態サブコーパス「雑誌」の母集団を「分野」で層別した際のタイトル数と冊数、ページ数について、表 2.3 に示す。

表 2.3: 生産実態サブコーパス「雑誌」の母集団（分野による層別）

分野	タイトル数		総冊数		総ページ数	
1. 総合	833	(66.16%)	38,383	(68.81%)	7,163,989	(68.79%)
2. 教育・学芸	163	(12.95%)	5,456	(9.78%)	983,224	(9.44%)
3. 政治・経済・商業	57	(4.53%)	3,168	(5.68%)	469,282	(4.51%)
4. 産業	12	(0.95%)	599	(1.07%)	115,172	(1.11%)
5. 工業	170	(13.50%)	7,101	(12.73%)	1,493,800	(14.34%)
6. 厚生・医療	24	(1.91%)	1,072	(1.92%)	189,488	(1.82%)
合計	1,259	(100.00%)	55,779	(100.00%)	10,414,955	(100.00%)

参照 雑誌を分野と発行年で層別した場合の冊数とページ数については、資料編の表 4.3 を参照（37 ページ）。

2.1.3 新聞の母集団の定義と層別

最後に、生産実態サブコーパス「新聞」の母集団の定義と、層別の方法について述べる。

「新聞」という形態には、全国紙、地方紙、スポーツ紙、専門紙、タウン紙など、幅広い様式が存在する。これらをすべて収録対象とすると、研究用途上あるいはコーパス構築の実務上、何らかの障害を生じさせる可能性が高い。そこで、『全国新聞ガイド』（社団法人日本新聞協会発行）において「全国紙」「ブロック紙」として記載されている日刊新聞」という条件で新聞タイトルを絞り込んだ。さらに、この条件ではカバーできない各地域の有力な地方紙も取り入れ、日本全国で発行されてい

る新聞を含むように調整した。その結果、2001年から2005年の間に発行された「新聞」の総体は図2.4のようになった。この結果を、生産実態サブコーパス「新聞」の母集団として定義した⁵。なお、ページ数に関しては、縮刷版のある全国紙4紙はランダムに取得した2003年の1週間分のページ数の平均値で求めた。その他は、直接新聞社へ電話で確認を行った。

タイトル数：16 タイトル（異なり）
冊数：49,625 冊
ページ数：1,198,189 ページ

図 2.4: 5 年間ににおける新聞のタイトル数・冊数・ページ数

新聞の母集団となる新聞 16 タイトルを、以下に示す。

全国紙：朝日新聞，毎日新聞，読売新聞，日本経済新聞，産経新聞

ブロック紙：北海道新聞，中日新聞，西日本新聞

地方紙：河北新報，新潟日報，京都新聞，神戸新聞，中国新聞，高知新聞，愛媛新聞，琉球新報

図 2.5: 新聞の母集団となるタイトル一覧

次に、この母集団を「紙種」「発行年」という基準で層別した。「紙種」は全国紙・ブロック紙・地方紙の別、および16種の新聞のタイトルとした。これに「発行年」の5分類を重ね合わせ、新聞全体を合計80の層に分割した。

新聞の母集団を「紙種」で層別した際の冊数とページ数の分布について、表2.4に示す。

表 2.4: 生産実態サブコーパス「新聞」の母集団（紙種による層別）

紙種	タイトル数	総冊数	総ページ数
全国紙	5 (31.25%)	15,950 (32.14%)	426,472 (35.59%)
ブロック紙	3 (18.75%)	9,570 (19.28%)	248,300 (20.72%)
地方紙	8 (50.00%)	24,105 (48.57%)	523,417 (43.68%)
合計	16 (100.00%)	49,625 (100.00%)	1,198,189 (100.00%)

参照 新聞をタイトル別に層別した場合の冊数とページ数については、資料編の表 4.4 を参照（38 ページ）。

以上が、BCCWJ 生産実態サブコーパスにおける母集団の定義および層別の結果である。

⁵ この場合の1冊は、1つの朝刊または夕刊を指す。

2.2 調査方法

次に、「現代日本語書き言葉の文字数調査」の調査方法について述べる。以下では、文字数を計測するページの抽出方法、計測対象とする文字要素、文字数計測の方法について述べる。

2.2.1 計測ページの抽出

まず、文字数を計測するページを抽出することを目的として、母集団を構成する各層を「判型（紙面のサイズ）」という基準によりさらに下位の層に分割した。判型の基準は、書籍の場合は J-BISC に記載されている「大きさ」（書籍の高さ）を集計し、「15cm 以下」から「31cm 以上」まで、および「記録なし」の計 18 種類に分類したものを用いた。雑誌の場合は『雑誌新聞総かたろぐ』に記載されている「判型」を用いた。これにより、例えば「NDC が 9 番台、高さ 21cm の書籍」「分野が 1 番（総合）、A4 判の雑誌」などの層ができる。新聞はほぼ定型なので、判型による層別は行わなかった。ここで得られる最下層に含まれるページの中から、複数のページを抽出することにした。

計測ページは、2003 年に発行された書籍・雑誌・新聞から抽出することにした。2003 年に発行された書籍・雑誌・新聞の総数は、以下ようになる。

書籍：65,719 冊，15,544,246 ページ
雑誌：909 タイトル（異なり），11,167 冊，2,095,217 ページ
新聞：16 タイトル（異なり），9,925 冊，239,638 ページ

図 2.6: 2003 年に発行された書籍・雑誌・新聞の総数

以下では、書籍・雑誌・新聞の別に、計測ページの抽出方法について述べる。

書籍 2003 年に発行された書籍は、合計 65,719 冊，15,544,246 ページである。これらについて、判型の種類ごとの冊数を、NDC 別に求めた。例えば、NDC が「0. 総記」に分類される書籍（2,248 冊）の判型ごとの冊数の分布は、表 2.5 のようになる。

参照 書籍の NDC と判型ごとの冊数については、資料編の表 4.5 を参照（39 ページ）。

ここで、冊数の累計が全体の 90%を超えたところまでの判型の種類（表 2.5 の場合は 18cm まで）を対象として、各判型に含まれる総冊数の約 5%をランダムに選び出した。さらに、選ばれた各冊の中から 5 ページをランダムに（もしくは等間隔に）抽出し、これらを書籍の計測ページとした。

以上のような手続きで抽出された冊数とページ数は、227 冊（2003 年に発行された書籍全体の 0.345%）、1,135 ページ（同 0.073 %）となった。NDC と判型ごとの冊数の分布を、表 2.6 に示す。

雑誌 2003 年に発行された雑誌は、異なりで 909 タイトル，合計 11,167 冊，2,095,217 ページである。これらについて、判型の種類ごとの冊数を、分野別に求めた。例えば、分野が「1. 総合」に分類される雑誌（600 冊）の判型ごとの分布は、表 2.7 のようになる。

表 2.5: 書籍 (NDC「0. 総記」) の判型ごとの冊数

判型	冊数	冊数累計	判型	冊数	冊数累計
21cm	464	464 (20.6%)	23cm	98	2,006 (89.2%)
19cm	464	928 (41.3%)	18cm	79	2,085 (92.7%)
20cm	301	1,229 (54.7%)	15cm	71	2,156 (95.9%)
24cm	270	1,499 (66.7%)	16cm	45	2,201 (97.9%)
26cm	179	1,678 (74.6%)	27cm	13	2,214 (98.5%)
22cm	115	1,793 (79.8%)	31cm	5	2,219 (98.7%)
30cm	115	1,908 (84.9%)	その他	29	2,248 (100.0%)

表 2.6: 書籍の計測対象として抽出された冊数

	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	計
0.	0	0	0	1	2	1	2	1	1	1	0	1	0	0	0	1	0	11
1.	1	3	0	1	4	2	2	2	0	0	0	0	0	0	0	0	0	15
2.	1	1	0	1	3	6	3	2	0	0	0	2	1	0	0	0	0	20
3.	1	0	0	2	10	7	9	15	0	0	0	4	0	0	0	3	0	51
4.	0	0	1	1	3	1	4	2	0	0	0	4	1	0	0	1	0	18
5.	1	1	0	1	2	2	4	2	0	1	0	3	1	0	1	3	0	22
6.	0	0	0	1	3	1	2	3	0	0	0	1	1	0	0	2	0	14
7.	1	1	0	1	4	4	4	1	0	0	0	2	0	0	1	1	1	21
8.	1	1	0	1	1	1	1	1	0	0	0	1	0	0	0	0	0	8
9.	10	5	0	5	8	12	2	5	0	0	0	0	0	0	0	0	0	47
計	16	12	1	15	40	37	33	34	1	2	0	18	4	0	2	11	1	227

行は NDC, 列は判型 (単位は cm) を表す。
判型のうち「15」は「15cm 以下」,「31」は「31cm 以上」を表す。

表 2.7: 雑誌 (分野「1. 総合」) の判型ごとの冊数

判型	冊数	冊数累計	判型	冊数	冊数累計
A4 変型判	342	342 (57.00%)	B6 判	5	582 (97.00%)
AB 判	94	436 (72.67%)	A6 判	3	585 (97.50%)
B5 判	77	513 (85.50%)	B4 変型判	3	588 (98.00%)
A4 判	31	544 (90.67%)	B6 変型判	3	591 (98.50%)
A5 判	27	571 (95.17%)	その他	9	600 (100.00%)
B5 変型判	6	577 (96.17%)			

参照 雑誌の分野と判型ごとの冊数については、資料編の表 4.6 を参照（40 ページ）。

ここで、冊数の累計が全体の 90%を超えたところまでの判型の種類（表 2.7 の場合は A4 判まで）を対象として、各判型に含まれる総冊数の約 5%をランダムに選び出した。さらに、選ばれた各冊の中から 5 ページをランダムに（もしくは等間隔に）抽出し、これらを雑誌の計測ページとした。

以上のような手続きで抽出された冊数とページ数は、53 冊（2003 年に発行された雑誌全体の 0.475%）、265 ページ（同 0.127%）となった。分野と判型ごとの冊数の分布を、表 2.8 に示す。

表 2.8: 雑誌の計測対象として抽出された冊数

分野	A4 変型判	A4 判	B5 判	AB 判	A5 判	菊倍判	その他	計
1. 総合	19	2	4	5	1	0	1	32
2. 教育・学芸	1	0	3	0	3	0	0	7
3. 政治・経済・商業	2	0	1	0	0	0	0	3
4. 産業	0	0	1	0	0	0	0	1
5. 工業	7	0	0	0	0	0	0	7
6. 厚生・医療	2	0	1	0	0	0	0	3
計	31	2	10	5	4	0	1	53

行は分野，列は判型を表す。

新聞 新聞の場合は、書籍・雑誌とは異なり、すべての紙種が「ブランクセット版」という判型で統一されている。そこで、全国紙のうち縮刷版のある「朝日」「毎日」「読売」「日経」を対象に文字数の計測を行い、その結果から母集団全体の文字数を推計することにした。そこで、全国 4 紙から、2003 年に発行された朝夕刊、各 1 日分を選び出し、1 冊全体を新聞の計測ページとした。1 日分を選び出す際、曜日のばらつきを考慮した⁶。

2003 年に発行された母集団に含まれる新聞は、異なりで 16 タイトル、9,925 冊、239,638 ページである。そのうち計測対象として抽出した冊数とページ数は、4 紙の朝夕刊 8 冊（全体の 0.081%）、合計 211 ページ（全体の 0.0880%）となった。

以上で、「現代日本語書き言葉の文字数調査」の計測対象ページを抽出することができた。

2.2.2 計測の対象とする文字要素

文字数の計測対象とするのは、印刷紙面上に現れる文字とした。文章内の論理構造の別（本文、図表、脚注、キャプション、ルビ、柱など）や、文字種の別（かな・カナ・漢字・記号・外国語・絵文字など）を問わず、印刷紙面上に現れた文字要素についてはすべてカウント対象とすることにした。

ただし、表紙・広告に現れる文字は一括して除外した。新聞の全面広告の場合は、柱の部分のみを計測対象とした。

⁶ ただし、曜日によっては例えば「株式欄」がないなど、面種の構成が異なることがある。そこで、選び出した 8 冊中、特に株式欄のない曜日の 2 冊分については、他の曜日の株式欄を計測ページとして補填した。

2.2.3 計測の方法

文字数の計測には、次の二通りの方法を採用した。

方法 1. 目視による計測：印刷紙面上にある文字を、人の目により数える方法。

方法 2. OCR による計測：印刷紙面を OCR ソフトで読み込み、自動文字認識により電子テキスト化して文字数を数える方法。

方法 1 では、計測ページとして選ばれたページをコピーし、そこに印刷されている文字数を目視により計測した⁷。方法 1 により計測したのは、書籍の一部と雑誌の一部、新聞のすべてである。方法 2 では、計測ページとして選ばれたページをコピーし、スキャナと OCR ソフトで自動的文字認識を行い、電子テキストとなった文字数を自動的にカウントした。方法 2 により計測したのは、書籍の一部と雑誌の一部である。

なお、方法 1 と方法 2 の違いが、計測結果にどの程度影響するかを確認するため、両者の結果を比較し、誤差を計測する作業を行った。まず、目視によって計測を行った書籍 15 冊、雑誌 10 冊について、OCR ソフトを用いて計測を行い、1 冊当たり 2 ページ分の結果を比較した。比較結果の抜粋を、表 2.9 に示す。「誤差」は、OCR による計測結果を目視による計測結果で割った値である。

表 2.9: 目視計測と OCR 計測の結果の誤差（一部）

書籍タイトル	計測ページ	目視計測	OCR 計測	誤差
世界遺産・大自然と秘境のミステリー （PHP 研究所）	116	381	395	103.67%
	223	426	429	100.70%
もっと知りたい坂本竜馬 （日本実業出版社）	152	791	794	100.38%
	191	743	742	99.87%
自衛隊はどのようにして生まれたか （学研）	65	687	690	100.44%
	176	620	620	100.00%

雑誌タイトル	計測ページ	目視計測	OCR 計測	誤差
Animage 2003 年 11 月号（徳間書店）	19	181	188	103.87%
	24	932	939	100.75%
HOBBY JAPAN 2003 年 9 月号（ホビージャパン）	134	3,338	3,159	94.64%
	157	760	819	107.76%
My Birthday 2003 年 11 月号（実業之日本社）	90	1,480	1,478	99.86%
	117	2,997	2,451	81.78%

書籍の 15 冊、30 ページ分について誤差（OCR 計測/目視計測）を平均したところ、102.22%となった。また、雑誌の 10 冊、20 ページ分について誤差を平均したところ、99.11%となった。特に図鑑などの場合、写真の上の色つき文字を OCR が認識できないなど、いくつかの問題があったが、全体の平均値の誤差が $\pm 3\%$ に収まることから、目視計測と OCR 計測の違いによる誤差は許容できる範囲であると見なした。

⁷ 計測の補助手段として電卓を用いた。

参照 目視による計測結果とOCRによる計測結果の誤差をまとめた表については、資料編の表4.7，表4.8を参照（41ページ）。

以上で、「現代日本語書き言葉の文字数調査」を実施するための準備が整った。計測対象となった冊数とページ数を、計測方法の別に、表2.10に示す。

表 2.10: 目視計測を実施した冊数と OCR 計測を実施した冊数

	書籍		雑誌		新聞	
目視計測	156 冊	780 ページ	28 冊	140 ページ	8 冊	211 ページ
OCR 計測	71 冊	355 ページ	25 冊	125 ページ	0 冊	0 ページ
合計	227 冊	1,135 ページ	53 冊	265 ページ	8 冊	211 ページ

第3章 「現代日本語書き言葉の文字数調査」の結果

本章では、母集団の各層から計測対象ページを抽出し、そこに現れる文字の数を計測した結果について示す。また、計測結果から、1 ページに含まれる文字の密度を表す「キャラクタ密度 (character density)」を層ごとに算出し、5 年分の総ページ数に掛け合わせることで、母集団全体に含まれる文字数を推計した結果を示す。さらに、その結果を用いて、生産実態サブコーパスの固定長サンプル 1,000 万語分のサンプル構成比を算出し、量的なコーパスデザインを完成させた結果について示す。

3.1 調査結果

3.1.1 書籍の文字数調査の結果

はじめに、書籍の文字数調査の結果を示す。

1 冊から 5 ページずつ抽出した計測ページに含まれる文字数を、目視または OCR によって計測した。結果を集計し、1 ページあたりに含まれる文字数の平均値を求めた。この平均値を、1 ページ中に含まれる文字の密度を表す数として、「キャラクタ密度」とした。書籍の文字数計測の結果を、表 3.1 に抜粋する。

表 3.1: 書籍の文字数計測結果 (一部)

NDC 判型	タイトル 出版社	頁 1 文字数	頁 2 文字数	頁 3 文字数	頁 4 文字数	頁 5 文字数	キャラクタ 密度
0. 総記 19cm	勝ち残る SE の絶対条件 (ナツメ社)	39 13	77 642	116 537	154 684	193 432	461.60
0. 総記 22cm	丸山眞男集 4 (岩波書店)	50 0	383 210	8 637	15 827	245 837	502.20
1. 哲学 19cm	神の発明 (講談社)	35 620	69 681	104 592	139 20	173 456	473.80
1. 哲学 22cm	哲学論集 (岩波書店)	120 726	122 460	142 435	369 517	277 264	480.40

参照 書籍 227 冊、1,135 ページの文字数計測の結果の詳細は、資料編の表 4.9、表 4.10、表 4.11、表 4.12、表 4.13 を参照 (43 ~ 47 ページ)。

計測対象となった227冊、1,135ページの文字数を計測することにより、各冊ごとにキャラクタ密度を求めた。ここで、キャラクタ密度の分布表として、11種類のNDCと18種類の判型による行列を作成し、各冊が該当するマスにキャラクタ密度を埋め込んだ。1つのマスに複数の冊が該当する場合は、キャラクタ密度の平均値を埋め込んだ。

これによって得られた判型ごとのキャラクタ密度の分布を集計し、回帰直線 $y=31.255x+379.5$ を得た¹。判型別のキャラクタ密度の分布と回帰直線を、図3.1に示す。

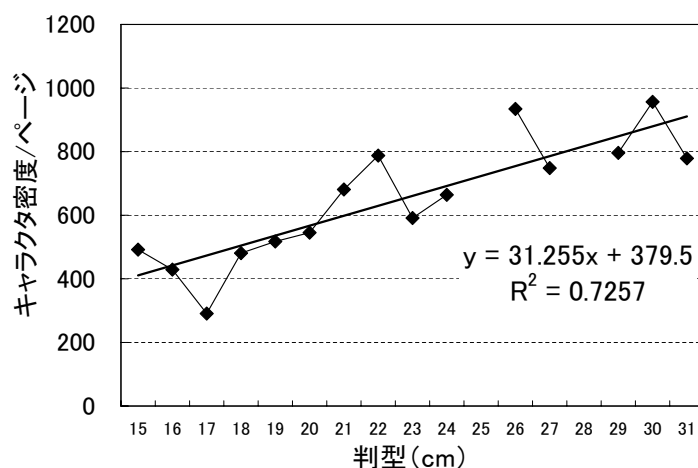


図 3.1: 書籍の判型別キャラクタ密度と回帰直線

この回帰直線を用いて、計測しなかったNDC・判型のマスについても、キャラクタ密度を計算した。これにより完成した行列を、書籍の「総文字数推計基準表」とした。完成した総文字数推計基準表を、表3.2に示す。斜体になっている数値は、回帰直線に基づいて算出された値であることを表す。

この「推計基準表」を、5年間におけるNDC・判型ごとのページ数に掛け合わせることで、書籍の5年間の総文字数を推計した。その結果、48,539,925,351文字という推計値を得た。表3.3に、NDCごとに総文字数を推計した結果を示す。

参照 書籍の文字調査のまとめとして、書籍の発行年、NDC、判型ごとに整理した冊数、ページ数、推計総文字数の分布を、資料編の表4.14、表4.15、表4.16に示す(48～52ページ)。

3.1.2 雑誌の文字数調査の結果

次に、雑誌の文字数調査の結果を示す。

書籍の場合と同様、1冊から5ページずつ抽出した計測ページに含まれる文字数を、目視またはOCRによって計測した。ただし、雑誌には広告が多く含まれることから、文字数の推計対象とすべきページ数は、図2.3で示した母集団の総ページ数(10,414,955ページ)よりも少なくなることになる。

¹ 判型とNDCごとのキャラクタ密度の分布についても集計したが、有効な回帰直線が得られなかったため、NDCごとの分布は用いなかった。

表 3.2: 書籍の総文字数推計基準表

NDC	15cm	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記	410.8	442.0	473.3	530.8	557.8	389.8	660.9	502.2	591.2
1. 哲学	544.0	418.5	473.3	542.2	497.0	529.6	743.9	655.8	660.8
2. 歴史	516.0	434.0	473.3	466.6	471.9	529.8	652.9	747.1	660.8
3. 社会科学	413.8	442.0	473.3	456.0	602.5	603.0	618.6	885.9	660.8
4. 自然科学	410.8	442.0	290.8	457.0	469.1	631.6	680.8	801.1	660.8
5. 技術工学	430.4	396.0	473.3	340.8	409.9	539.0	672.8	815.5	660.8
6. 産業	410.8	442.0	473.3	558.8	513.9	511.6	448.3	1192.7	660.8
7. 芸術	503.6	492.8	473.3	451.2	573.4	654.5	667.1	638.8	660.8
8. 言語	600.6	345.0	473.3	522.2	632.0	561.6	910.4	1052.0	660.8
9. 文学	435.2	487.2	473.3	482.8	447.8	501.4	753.3	585.3	660.8
null	410.8	442.0	473.3	504.5	535.8	567.0	598.3	629.5	660.8

NDC	24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm	null
0. 総記	467.6	723.3	1026.4	785.8	817.1	848.3	569.4	910.8	629.9
1. 哲学	692.1	723.3	754.6	785.8	817.1	848.3	879.6	910.8	675.1
2. 歴史	692.1	723.3	1234.6	851.6	817.1	848.3	879.6	910.8	700.6
3. 社会科学	692.1	723.3	1272.6	785.8	817.1	848.3	1674.1	910.8	757.6
4. 自然科学	692.1	723.3	1159.7	473.6	817.1	848.3	580.4	910.8	650.0
5. 技術工学	860.8	723.3	702.4	544.8	817.1	636.2	962.5	910.8	641.0
6. 産業	692.1	723.3	1254.6	1123.2	817.1	848.3	441.5	910.8	707.2
7. 芸術	692.1	723.3	318.5	785.8	817.1	955.8	1510.8	778.8	688.1
8. 言語	692.1	723.3	506.6	785.8	817.1	848.3	879.6	910.8	701.3
9. 文学	692.1	723.3	754.6	785.8	817.1	848.3	879.6	910.8	661.1
null	692.1	723.3	754.6	785.8	817.1	848.3	879.6	910.8	660.8

表 3.3: 書籍の文字数調査のまとめ — NDC 別の推計総文字数

NDC	総文字数	構成比	NDC	総文字数	構成比
0. 総記	1,636,414,548	3.371%	6. 産業	2,196,387,437	4.525%
1. 哲学	2,597,610,813	5.351%	7. 芸術	3,258,432,447	6.713%
2. 歴史	4,301,204,340	8.861%	8. 言語	888,800,128	1.831%
3. 社会科学	12,408,321,943	25.563%	9. 文学	9,341,275,486	19.245%
4. 自然科学	5,069,594,034	10.444%	n. 記録なし	2,225,954,208	4.586%
5. 技術工学	4,615,929,967	9.510%	合計	48,539,925,351	100.00%

そこで、対象とした冊ごとに広告を除いたページ数を数え、計測対象になり得るページ数を求めた。次に、計測した5ページ分の文字数の平均値を計測対象ページ数に掛け合わせ、1冊に含まれる総文字数を推計した。この1冊中の総文字数を、最初に母集団のページ数を集計した『雑誌新聞総かたろぐ』に掲載されている当該タイトルのページ数で割り、その冊の「キャラクタ密度」とした。計測結果を、表3.4に抜粋する。

表 3.4: 雑誌の文字数推計結果（一部）

分野 判型	タイトル 出版社	頁1 文字数	頁2 文字数	頁3 文字数	頁4 文字数	頁5 文字数	1	2
総合 AB判	My Birthday (実業之日本社)	117 2,997	90 1,480	69 5,774	3 2,009	138 3,210	143 222	442,442.0 1,993.0
総合 A4変	animage (徳間書店)	37 2,376	24 932	19 181	45 4,460	16 450	143 220	240,211.4 1,091.9
総合 B5	HOBBY JAPAN (ホビージャパン)	200 303	157 760	134 3,338	682 745	17 726	219 376	257,193.6 684.0
総合 A4変	AERA (朝日新聞社)	15 2,193	30 2,372	44 2,555	61 1,125	75 2,115	71 80	146,473.2 1,830.9

1の列は、上段が計測対象ページ数、下段が『雑誌新聞総かたろぐ』掲載のページ数を表す。

2の列は、上段が1冊あたりの総文字数、下段がキャラクタ密度を表す。

参照 雑誌53冊、265ページの文字数計測の結果の詳細は、資料編の表4.17を参照(54ページ)。

計測対象となった53冊、265ページの文字数を計測することにより、各冊ごとにキャラクタ密度を求めた。ここで、6種類の分野と5種類にまとめ上げた判型による行列を作成し、各冊が該当するマスにキャラクタ密度を埋め込んだ。1つのマスに複数の冊が該当する場合は、キャラクタ密度の平均値を埋め込んだ。また、実際に計測しなかった分野・判型については、他の分野で計測したその判型のキャラクタ密度の平均値を割り当てた。この結果を、雑誌の「総文字数推計基準表」とした。完成した総文字数推計基準表を、表3.5に示す。斜体になっている数値は、他の分野で計測した平均値に基づいて算出された値であることを表す。

表 3.5: 雑誌の総文字数推計基準表

分野	A4系	A5系	AB系	B4系	B5系
1. 総合	1022.1	1031.2	1413.7	831.1	807.5
2. 教育・学芸	765.4	884.8	1413.7	831.1	878.1
3. 政治・経済・商業	1014.2	921.4	1413.7	831.1	798.2
4. 産業	1093.2	921.4	1413.7	831.1	506.6
5. 工業	973.2	921.4	1413.7	831.1	767.9
6. 厚生・医療	1273.3	921.4	1413.7	831.1	510.0

この「推計基準表」を、5年間ににおける分野・判型ごとのページ数に掛け合わせ、雑誌の5年間の総文字数を推計した。その結果、10,515,681,636文字という推計値を得た。表3.6に、分野ごとの推計結果を示す。

表 3.6: 雑誌の文字数調査のまとめ — 分野別の推計総文字数

分野	総文字数	構成比
1. 総合	7,421,447,806	70.575%
2. 教育・学芸	877,875,592	8.348%
3. 政治・経済・商業	456,459,405	4.341%
4. 産業	110,640,958	1.052%
5. 工業	1,468,293,360	13.963%
6. 厚生・医療	180,964,513	1.721%
合計	10,515,681,636	100.00%

参照 雑誌の文字調査のまとめとして、雑誌の発行年、分野、判型ごとに整理した冊数、ページ数、推計文字数の分布を、資料編の表 4.18、表 4.19、表 4.20 に示す（56～60 ページ）。

3.1.3 文字数調査の結果 (3) — 新聞

最後に、新聞の文字数調査の結果について述べる。

まず、全国紙 4 紙の朝夕刊、計 8 冊、全 211 ページを対象として、目視により文字数を計測した。

参照 特に新聞の場合は、印刷紙面を便宜的に複数の区画に分割し、各区画に含まれる文字数を数え上げて最後に合計するという作業を行った。実際の文字数計測に用いた新聞の作業コピーの例を、資料編の図 4.1、図 4.2 に挙げる（62～63 ページ）。

次に、実際に計測した 1 冊の紙面構成を分析し、面種（いわゆる社会面・政治面など）ごとの面積と 1cm^2 あたりに含まれる文字数を算出した。算出の結果を、表 3.7 に抜粋する。

参照 新聞（朝刊）の文字数計測の結果の詳細は、資料編の表 4.21、表 4.22、表 4.23、表 4.24 を参照。（64～67 ページ）

さらに、計測した冊の発行日以外の曜日の冊を 1 週間分ランダムに抽出し、それぞれの面種ごとに面積を計測して、4 紙の 1 週間分の紙面構成と面積を明らかにした。面種ごとに求めた 1cm^2 あたりの文字数を、面種ごとの総面積に掛け合わせるにより、新聞 1 週間分に含まれる総文字数を推計した。例を表 3.8 に示す。

参照 実際に文字数計測を行った冊の日付と、面積を計測した日付の一覧は、資料編の表 4.25 を参照（68 ページ）。また、4 紙の面種をどのようにまとめ上げたかについては、資料編の表 4.26 を参照（69 ページ）。さらに、4 紙 1 週間分の文字数推計の結果は、資料編の表 4.27、表 4.28、表 4.29、表 4.30 を参照（70～73 ページ）。

表 3.7: 朝日新聞朝刊の面種ごとの面積と計測した文字数（一部）

面	面種	文字数	単純面積	文字数/cm ²
1	一面	5,342	370	14.44
2	総合	5,220	305.25	17.10
3	総合	5,508	305.25	18.04
4	国際	4,974	305.25	16.29
5	経済	3,810	249.75	15.26
6	柱情報なし（「教師力」）	5,427	305.25	17.78
7	全面広告	43	0	—
8	全面広告	43	0	—
9	柱情報なし（朝日俳壇・歌壇）	3,961	249.75	15.86
10	オピニオン	5,738	305.25	18.80

表 3.8: 朝日新聞 1 週間分の面種ごとの面積と推計総文字数（一部）

面種	面積	推計総文字数	面種	面積	推計総文字数
一面	2,590	37,394	家庭	3,608	51,337
総合	6,475	113,782	ラテ欄	6,910	166,774
国際	3,478	56,673	地域	4,283	78,594
経済	4,810	73,378	社会	7,039	115,420
特集	4,227	64,935	証券	3,506	291,912
オピニオン	4,209	73,972	合計	58,858	1,271,618
スポーツ	7,724	147,446			

「面積」の単位は cm² である。

以上の手続きで得られた1週間分の総文字数を平均して1日分の平均文字数を算出し、さらに朝刊は355日分、夕刊は283日分を掛け合わせるにより、1年間の総文字数を算出した。この1年間の総文字数を総ページ数で割ることにより、4紙の朝夕刊それぞれについてキャラクタ密度を算出した。ここでは朝日新聞朝刊のキャラクタ密度の算出結果について、表3.9に示す。

表 3.9: 朝日新聞朝刊 1 週間分の文字数・ページ数と、キャラクタ密度の算出

	月	火	水	木	金	土	日	1 日平均
文字数	124,764	194,589	211,046	183,908	193,373	221,355	142,579	181,659.14
ページ数	36	36	40	36	32	36	40	36.57

	1 日平均	1 年分	キャラクタ密度
文字数	181,659.14	64,488,995.71	4,967.24
ページ数	36.57	12,982.86	

参照 4 紙の朝夕刊それぞれについて 1 週間分の文字数・ページ数と、キャラクタ密度の算出した結果の一覧を、資料編の表 4.31 に挙げる（74 ページ）。

さらに、計測しなかった新聞タイトルについては、朝日・毎日・読売のキャラクタ密度の平均値をキャラクタ密度とした²。このキャラクタ密度を5年間の総ページ数と掛け合わせ、新聞の5年間の総文字数を推計した。その結果、約 64.2 億文字という推計値を得た。ここでは、新聞タイトルごとの推計結果を表 3.10 に示す。

参照 16 紙の朝夕刊それぞれについて、5 年間の総文字数・キャラクタ密度を算出した結果の一覧を、資料編の表 4.32 に挙げる（75 ページ）。

以上で、書籍・雑誌・新聞に関する5年間の総文字数の推計値が出揃った。

3.2 調査のまとめ

「現代日本語書き言葉の文字数調査」の実施により、定義された母集団（2001 年から 2005 年に生産された書籍・雑誌・新聞）に含まれる総文字数の推計値として、書籍が約 485.4 億文字、雑誌が約 105.2 億文字、新聞が約 64.2 億文字、合計すると約 654.7 億字という結果を得た。

調査結果全体をまとめて各層の比を算出した結果を、表 3.11 に示す。また、書籍・雑誌・新聞のそれぞれについて、発行年と NDC・分野・新聞タイトルで層別した構成比を円グラフにしたものを、図 3.2、図 3.3、図 3.4 に示す。

² 日本経済新聞は株価欄が多く、キャラクタ密度が他と比べて多かったため、キャラクタ密度の平均値を求める際には除外した。

表 3.10: 新聞の文字数調査のまとめ — 新聞タイトル別の推計総文字数

区分	新聞タイトル	総文字数	構成比
全国紙	朝日新聞	461,946,416	7.200%
	毎日新聞	399,078,164	6.220%
	読売新聞	457,239,476	7.126%
	日本経済新聞	688,108,023	10.725%
	産経新聞	411,250,382	6.410%
ブロック紙	北海道新聞	420,730,356	6.557%
	中日新聞	455,131,442	7.094%
	西日本新聞	420,730,356	6.557%
地方紙	河北新報	368,688,876	5.746%
	新潟日報	351,048,482	5.471%
	京都新聞	357,758,315	5.576%
	神戸新聞	377,170,042	5.879%
	中国新聞	351,048,482	5.471%
	高知新聞	298,127,299	4.647%
	愛媛新聞	246,965,522	3.849%
	琉球新報	351,048,482	5.471%
合計		6,416,070,114	100.00%

表 3.11: 2001 年から 2005 年の書籍・雑誌・新聞の推計総文字数

層		総文字数	構成比	層		総文字数	構成比
書籍	0. 総記	1,636,414,548	2.50%	雑誌	1. 総合	7,421,447,806	11.34%
	1. 哲学	2,597,610,813	3.97%		2. 教育	877,875,592	1.34%
	2. 歴史	4,301,204,340	6.57%		3. 政治	456,459,405	0.70%
	3. 社会科学	12,408,321,943	18.95%		4. 産業	110,640,958	0.17%
	4. 自然科学	5,069,594,034	7.74%		5. 工業	1,468,293,360	2.24%
	5. 技術工学	4,615,929,967	7.05%		6. 厚生	180,964,513	0.28%
	6. 産業	2,196,387,437	3.35%	新聞	全国紙	2,417,622,461	3.69%
	7. 芸術	3,258,432,447	4.98%		ブロック紙	1,296,592,154	1.98%
	8. 言語	888,800,128	1.36%		地方紙	2,701,855,499	4.13%
	9. 文学	9,341,275,486	14.27%	合計		65,471,677,100	100%
	n. 記録なし	2,225,954,208	3.40%				

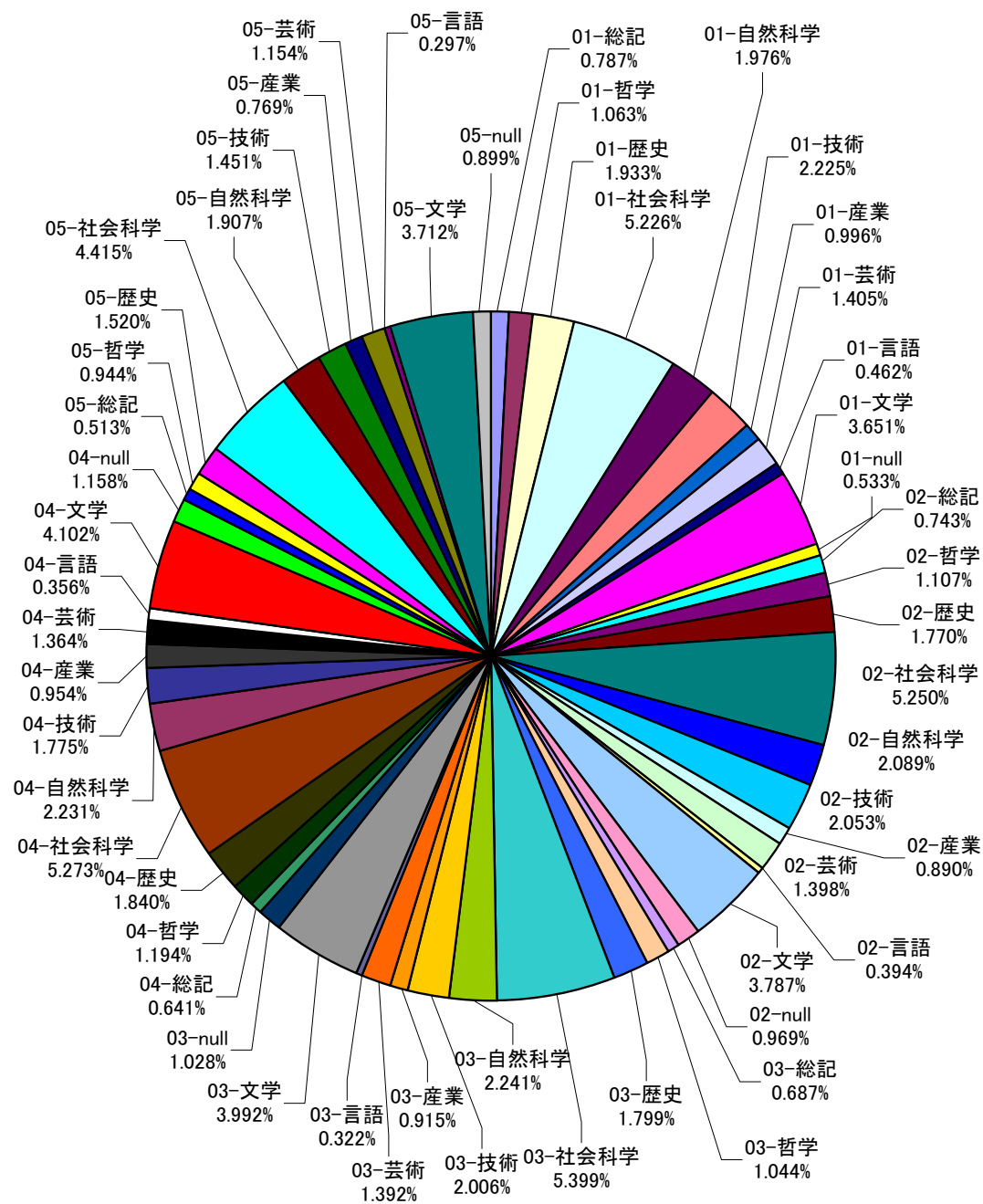


図 3.2: 書籍の総文字数構成比

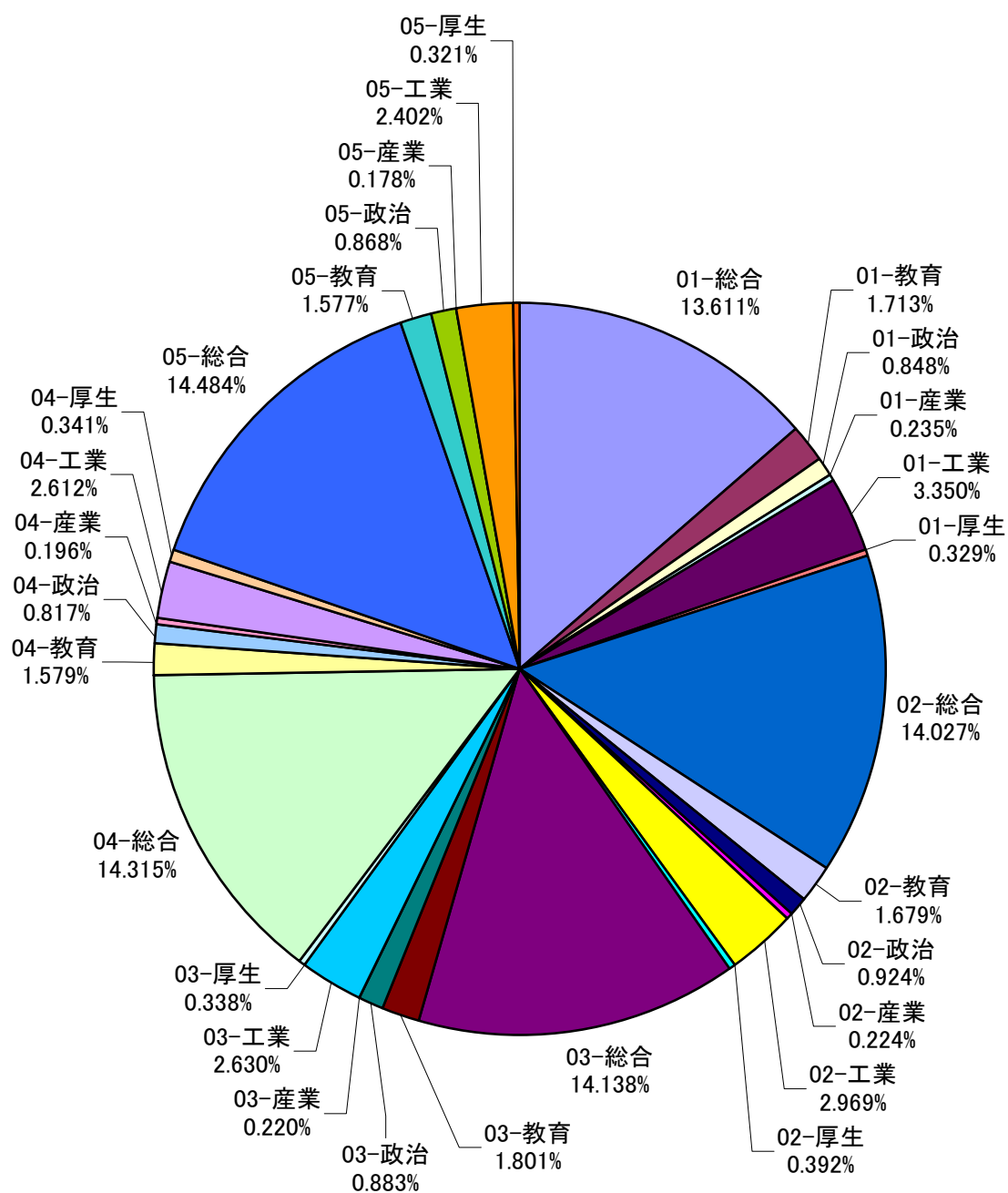


図 3.3: 雑誌の総文字数構成比

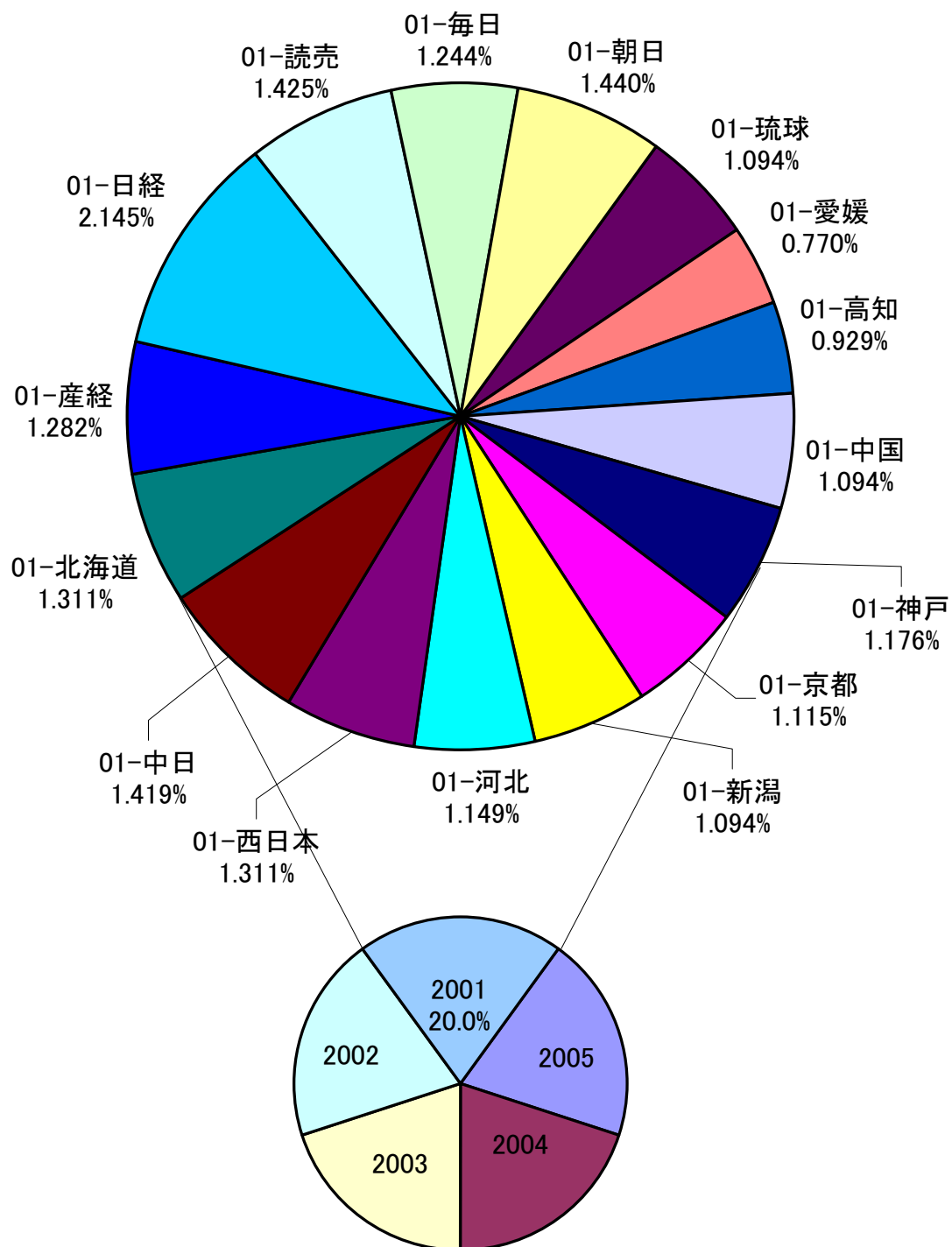


図 3.4: 新聞の総文字数構成比

3.3 サンプル構成比の算出

前節までで示した「現代日本語書き言葉の文字数調査」によって、母集団として定義した書籍・雑誌・新聞の総文字数を推計することができた。以下では、この文字数調査の結果を受けて、生産実態サブコーパスの固定長サンプル 1,000 万語分の構成比をどのように算出したかについて示す。

3.3.1 サンプル構成比の求め方

「現代日本語書き言葉の文字数調査」によって得られた、書籍・雑誌・新聞の推計総文字数は、以下のようなものであった。

表 3.12: 2001 年から 2005 年の書籍・雑誌・新聞の推計総文字数 (2)

	書籍	雑誌	新聞	合計
推計総文字数	48,539,925,351 (74.14%)	10,515,681,636 (16.06%)	6,416,070,114 (9.80%)	65,471,677,100 (100.00%)

ここで得られた比は、1.2 節で述べた設計方針に従って、生産実態サブコーパスに含まれる固定長サンプル 1,000 万語の内訳として採用される。すなわち、書籍からは 741.4 万語、雑誌からは 160.6 万語、新聞からは 98 万語をそれぞれ抽出するように、計画を立てればよい。書籍の 741.4 万語の内訳は、さらにその下位にある 55 の層（発行年、NDC）に含まれる総文字数の比によって決められることになる。雑誌・新聞も、それぞれ 30・80 の下位の層ごとに比が算出され、全体が構成される。

このように、推計した総文字数の比を 1,000 万語分のサンプル構成比として採用することにより、書き言葉の生産力を「文字数」という絶対的な量によって捉え、その分布のありさまを忠実に反映するサブコーパスを構築することができるわけである。その結果得られる 1,000 万語の内訳図を、図 3.5 に示す。さらに言えば、図 3.5 は、図 3.2、図 3.3、図 3.4 で全容を示した下位の層に分割され、各層から抽出される語数が決められることになる。

3.3.2 必要サンプル数の算出

固定長サンプル 1,000 万語分の構成比が以上のように決められたとして、各層から実際にいくつのサンプルが必要になるかも試算しておかなければならない。

例えば、図 3.2 で 2.50% という構成比を持つ書籍の「0. 総記」という層からは、25 万語を抽出することになる。1.2 節で述べたように、1 語が平均 1.7 文字で構成されたと考えると、この場合、1,000 文字の固定長サンプルを 425 サンプルを取得すれば、目標語数である 25 万語が達成される見込みとなる。また、可変長サンプルも固定長サンプルと同時に取られるので、書籍の可変長サンプルが 1 サンプルあたり平均 3,900 文字とすれば、約 975,000 語分の可変長サンプルが作られることになる。

そこで、この構成比に従って書籍・雑誌・新聞から抽出する語数を求め、必要な固定長サンプルの数を算出した。また、そこから得られる可変長サンプルの合計語数を見積もった。結果を表 3.13 に示す。これで、生産実態サブコーパス全体の量的なコーパスデザインが完成した。

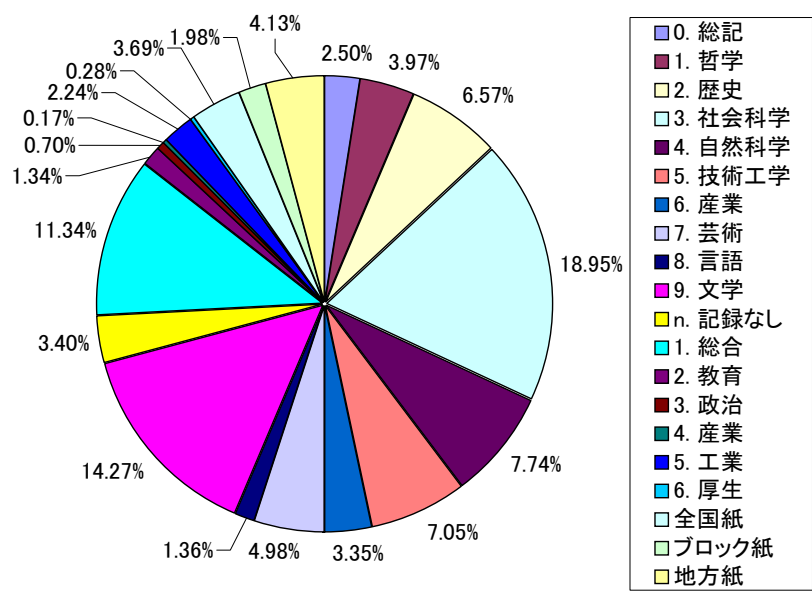


図 3.5: 固定長サンプル (1,000 万語分) の構成比

表 3.13: 生産実態サブコーパス全体のサンプル構成比

	総文字数	構成比	固定長 合計語数	必要サンプル数	可変長 合計語数
書籍	485.40 億文字	74.14%	741.4 万語	12,604 サンプル	2891.5 万語
雑誌	105.16 億文字	16.06%	160.6 万語	2,730 サンプル	481.8 万語
新聞	64.16 億文字	9.80%	98.0 万語	1,666 サンプル	98.0 万語
合計	654.72 億文字	100.00%	1,000 万語	17,000 サンプル	3471.3 万語

以上，本報告書では，BCCWJ の設計（コーパスデザイン）の一環として実施した「現代日本語書き言葉の文字数調査」について，その方法論と総文字数推計の結果を示した。さらに，推計された総文字数の比を，生産実態サブコーパスの固定長サンプル 1,000 万語分のサンプル構成比として反映させる手順について述べた。

参考文献

- 丸山岳彦・柏野和佳子・稲益佐知子・秋元祐哉・吉田谷幸宏・山崎誠 (2007) 書き言葉の構造を捉える ―書き言葉の多様な構造とサンプリング手法―. 『言語処理学会 第13回年次大会 発表論文集』. 言語処理学会
- 丸山岳彦・柏野和佳子・山崎誠・佐野大樹・秋元祐哉・稲益佐知子・吉田谷幸宏 (2007) 「現代日本語書き言葉均衡コーパス」におけるサンプリングの概要. 『特定領域「日本語コーパス」平成18年度公開ワークショップ（研究成果報告会）予稿集』. 国立国語研究所
- 山崎誠 (2007) 『現代日本語書き言葉均衡コーパス』の基本設計について. 『特定領域「日本語コーパス」平成18年度公開ワークショップ（研究成果報告会）予稿集』. 国立国語研究所

第II部

資料編

第4章 「現代日本語書き言葉の文字数調査」 資料集

表 4.1: 生産実態サブコーパス「書籍」の母集団（発行年・NDC による層別）

⇒ 参照元 2.1.1 節

	NDC	総冊数	総ページ数
2001	0. 総記	2,503 (0.79%)	667,715 (0.89%)
	1. 哲学	3,474 (1.10%)	900,008 (1.20%)
	2. 歴史	5,206 (1.64%)	1,391,806 (1.86%)
	3. 社会科学	12,813 (4.04%)	3,279,637 (4.38%)
	4. 自然科学	5,469 (1.72%)	1,295,428 (1.73%)
	5. 技術工学	7,211 (2.27%)	1,539,416 (2.05%)
	6. 産業	3,278 (1.03%)	709,456 (0.95%)
	7. 芸術	5,298 (1.67%)	1,083,312 (1.45%)
	8. 言語	1,317 (0.42%)	304,599 (0.41%)
	9. 文学	13,488 (4.25%)	3,558,890 (4.75%)
	n. 記録なし	2,914 (0.92%)	400,720 (0.53%)
2002	0. 総記	2,346 (0.74%)	633,884 (0.85%)
	1. 哲学	3,635 (1.15%)	929,325 (1.24%)
	2. 歴史	4,950 (1.56%)	1,287,397 (1.72%)
	3. 社会科学	12,938 (4.08%)	3,313,592 (4.42%)
	4. 自然科学	5,733 (1.81%)	1,357,459 (1.81%)
	5. 技術工学	6,692 (2.11%)	1,424,416 (1.90%)
	6. 産業	3,083 (0.97%)	659,593 (0.88%)
	7. 芸術	5,199 (1.64%)	1,069,813 (1.43%)
	8. 言語	1,143 (0.36%)	260,842 (0.35%)
	9. 文学	14,217 (4.48%)	3,710,845 (4.95%)
	n. 記録なし	4,417 (1.39%)	666,743 (0.89%)
2003	0. 総記	2,248 (0.71%)	584,306 (0.78%)
	1. 哲学	3,616 (1.14%)	887,380 (1.18%)
	2. 歴史	5,121 (1.61%)	1,310,844 (1.75%)
	3. 社会科学	13,292 (4.19%)	3,395,390 (4.53%)
	4. 自然科学	6,041 (1.90%)	1,446,005 (1.93%)
	5. 技術工学	6,424 (2.03%)	1,406,853 (1.88%)
	6. 産業	3,152 (0.99%)	674,493 (0.90%)
	7. 芸術	5,262 (1.66%)	1,067,470 (1.42%)
	8. 言語	942 (0.30%)	214,818 (0.29%)
	9. 文学	15,192 (4.79%)	3,905,502 (5.21%)
	n. 記録なし	4,428 (1.40%)	651,185 (0.87%)
2004	0. 総記	2,185 (0.69%)	540,869 (0.72%)
	1. 哲学	3,993 (1.26%)	1,007,377 (1.34%)
	2. 歴史	5,137 (1.62%)	1,338,088 (1.79%)
	3. 社会科学	12,924 (4.08%)	3,275,960 (4.37%)
	4. 自然科学	6,212 (1.96%)	1,451,958 (1.94%)
	5. 技術工学	6,012 (1.90%)	1,263,110 (1.69%)
	6. 産業	3,181 (1.00%)	690,668 (0.92%)
	7. 芸術	5,200 (1.64%)	1,045,430 (1.40%)
	8. 言語	974 (0.31%)	225,671 (0.30%)
	9. 文学	16,011 (5.05%)	4,034,730 (5.39%)
	n. 記録なし	4,897 (1.54%)	735,215 (0.98%)
2005	0. 総記	1,850 (0.58%)	433,019 (0.58%)
	1. 哲学	3,349 (1.06%)	805,239 (1.07%)
	2. 歴史	4,210 (1.33%)	1,121,037 (1.50%)
	3. 社会科学	11,019 (3.47%)	2,794,537 (3.73%)
	4. 自然科学	5,290 (1.67%)	1,221,108 (1.63%)
	5. 技術工学	5,038 (1.59%)	1,047,540 (1.40%)
	6. 産業	2,638 (0.83%)	564,103 (0.75%)
	7. 芸術	4,428 (1.40%)	887,506 (1.18%)
	8. 言語	835 (0.26%)	190,910 (0.25%)
	9. 文学	14,808 (4.67%)	3,678,311 (4.91%)
	n. 記録なし	3,884 (1.22%)	569,992 (0.76%)
合計		317,117 (100%)	74,911,520 (100%)

表 4.2: 2001 年から 2005 年の各年における雑誌協会加盟社のリスト

⇒ 参照元 2.1.2 節

出版社名	2001	2002	2003	2004	2005
1 秋田書店	○	○	○	○	○
2 朝日新聞社	○	○	○	○	○
3 アシエット婦人画報	○	○	○	○	○
4 アスキー	○	○	○	○	○
5 家の光協会	○	○	○	○	○
6 インプレス	○	○	○	○	○
7 潮出版社	○	○	○	○	○
8 SSコミュニケーションズ	○	○	○	○	○
9 エンターブレイン	○	○	○	○	○
10 旺文社	○	○	○	○	○
11 オークラ出版	○	○	○	○	○
12 オレンジページ	○	○	○	○	○
13 学習研究社	○	○	○	○	○
14 角川書店	○	○	○	○	○
15 河出書房新社	○	○	○	○	○
16 行研	○	○	○	-	-
17 共同通信社	○	-	-	-	-
18 近代映画社	○	○	○	○	○
19 暮しの手帖社	○	○	○	○	○
20 経済界	○	○	○	○	○
21 芸術生活社	○	○	○	○	○
22 京阪神エルマガジン	-	-	○	○	○
23 芸文社	○	○	○	○	○
24 弘済出版社	○	-	-	-	-
25 廣済堂出版	○	○	○	-	-
26 講談社	○	○	○	○	○
27 交通新聞社	-	○	○	○	○
28 交通タイムス社	○	○	○	○	○
29 光文社	○	○	○	○	○
30 ゴルフダイジェスト社	○	○	○	○	○
31 財界研究所	○	○	○	○	○
32 三栄書房	○	○	○	○	○
33 実業之日本社	○	○	○	○	○
34 シティ出版	○	○	○	○	○
35 集英社	○	○	○	○	○
36 主婦と生活社	○	○	○	○	○
37 主婦の友社	○	○	○	○	○
38 小学館	○	○	○	○	○
39 祥伝社	○	○	○	○	○
40 少年画報社	○	○	○	○	○
41 シンコーミュージック	○	○	○	○	○
42 新潮社	○	○	○	○	○
43 スキージャーナル	○	○	○	○	○
44 スクウェアエニックス	-	○	○	○	○
45 青春出版社	○	○	○	○	○
46 誠文堂新光社	○	○	○	○	○
47 世界文化社	○	○	○	○	○
48 ソニーマガジンズ	○	○	○	○	○
49 ダイヤモンド社	○	○	○	○	○
50 中央公論社	○	○	○	○	○
51 ディーアンドエー	-	○	○	○	○
出版社名	2001	2002	2003	2004	2005
52 電波新聞社	○	○	○	○	○
53 東京ニュース通信社	○	○	○	○	○
54 東洋経済新報社	○	○	○	○	○
55 徳間書店	○	○	○	○	○
56 日刊現代	○	○	○	○	○
57 日刊スポーツ出版社	○	○	○	○	○
58 日経BP	○	○	○	○	○
59 日経ホーム	○	○	○	○	○
60 日本放送出版協会	○	○	○	○	○
61 日本ジャーナル出版	○	○	○	○	○
62 日本スポーツ企画出版	○	○	○	○	○
63 日本スポーツ出版社	○	○	○	○	○
64 日本文芸社	○	○	○	○	○
65 日本ウォーク社	○	-	-	-	-
66 ニュートンプレス	○	○	○	○	○
67 白泉社	○	○	○	○	○
68 ひあ	○	○	○	○	○
69 PHP研究所	○	○	○	○	○
70 ビクターエンタテインメント	○	○	-	-	-
71 日之出出版	○	○	○	○	○
72 婦人生活社	○	○	-	-	-
73 婦人之友社	○	○	○	○	○
74 扶桑社	○	○	○	○	○
75 双葉社	○	○	○	○	○
76 プレック社	○	○	○	○	○
77 プレジデント社	○	○	○	○	○
78 ぶんか社	○	○	○	○	○
79 文化出版局	○	○	○	○	○
80 文藝春秋	○	○	○	○	○
81 平凡社	○	○	-	-	-
82 ベースボールマガジン	○	○	○	○	○
83 ベストセラーズ	-	○	○	○	○
84 ベネッセコーポレーション	-	-	-	○	○
85 ベルシステム24	-	-	-	○	○
86 芳文社	○	○	○	○	○
87 ホビージャパン	-	-	○	○	○
88 ポリスマガジン	-	-	-	○	○
89 毎日コミュニケーションズ	○	○	○	○	○
90 毎日新聞社	○	○	○	○	○
91 マガジンハウス	○	○	○	○	○
92 マガジンマガジン	-	-	-	-	○
93 マキノ出版	○	○	○	○	○
94 マリン企画	○	○	○	○	○
95 メディアワークス	○	○	○	○	○
96 八重洲出版	○	○	○	○	○
97 山と溪谷社	○	○	○	○	○
98 読売新聞	○	○	○	○	○
99 リイド社	○	○	○	○	○
100 陸上競技社	○	○	○	○	○
101 立風書房	○	○	-	-	-
102 ワニマガジン社	○	○	○	○	○

「○」はその年に加盟していたことを、「-」は加盟していなかったことを表す。

表 4.3: 生産実態サブコーパス「雑誌」の母集団（発行年・分野による層別）

⇒ 参照元 2.1.2 節

	分野	タイトル数	冊数	ページ数
2001	1.総合	594 (12.99%)	7,585 (13.60%)	1,371,908 (13.17%)
	2.教育・学芸	126 (2.75%)	1,100 (1.97%)	200,884 (1.93%)
	3.政治・経済・商業	46 (1.01%)	656 (1.18%)	92,560 (0.89%)
	4.産業	8 (0.17%)	125 (0.22%)	25,116 (0.24%)
	5.工業	136 (2.97%)	1,701 (3.05%)	357,048 (3.43%)
	6.厚生・医療	20 (0.44%)	203 (0.36%)	34,272 (0.33%)
2002	1.総合	595 (13.01%)	7,651 (13.72%)	1,422,765 (13.66%)
	2.教育・学芸	130 (2.84%)	1,148 (2.06%)	197,584 (1.90%)
	3.政治・経済・商業	48 (1.05%)	692 (1.24%)	98,850 (0.95%)
	4.産業	8 (0.17%)	117 (0.21%)	24,636 (0.24%)
	5.工業	124 (2.71%)	1,514 (2.71%)	316,861 (3.04%)
	6.厚生・医療	22 (0.48%)	239 (0.43%)	42,960 (0.41%)
2003	1.総合	600 (13.12%)	7,668 (13.75%)	1,441,592 (13.84%)
	2.教育・学芸	129 (2.82%)	1,179 (2.11%)	212,468 (2.04%)
	3.政治・経済・商業	42 (0.92%)	630 (1.13%)	95,508 (0.92%)
	4.産業	9 (0.20%)	131 (0.23%)	24,180 (0.23%)
	5.工業	109 (2.38%)	1,349 (2.42%)	284,317 (2.73%)
	6.厚生・医療	20 (0.44%)	210 (0.38%)	37,152 (0.36%)
2004	1.総合	607 (13.27%)	7,692 (13.79%)	1,452,792 (13.95%)
	2.教育・学芸	122 (2.67%)	1,005 (1.80%)	186,446 (1.79%)
	3.政治・経済・商業	38 (0.83%)	593 (1.06%)	88,260 (0.85%)
	4.産業	8 (0.17%)	119 (0.21%)	21,580 (0.21%)
	5.工業	108 (2.36%)	1,324 (2.37%)	279,794 (2.69%)
	6.厚生・医療	21 (0.46%)	214 (0.38%)	37,520 (0.36%)
2005	1.総合	619 (13.53%)	7,789 (13.96%)	1,474,932 (14.16%)
	2.教育・学芸	117 (2.56%)	1,024 (1.84%)	185,842 (1.78%)
	3.政治・経済・商業	40 (0.87%)	597 (1.07%)	94,104 (0.90%)
	4.産業	7 (0.15%)	107 (0.19%)	19,660 (0.19%)
	5.工業	100 (2.19%)	1,213 (2.17%)	255,780 (2.46%)
	6.厚生・医療	21 (0.46%)	206 (0.37%)	37,584 (0.36%)
合計		4,574 (100%)	55,779 (100%)	10,414,955 (100%)

タイトル数の合計は、述べてある。

表 4.4: 生産実態サブコーパス「新聞」の母集団（紙種・新聞タイトルによる層別）

⇒ 参照元 2.1.3 節

紙種	新聞タイトル		冊数	ページ数 ／冊	総ページ数
全国	朝日新聞	朝刊	1,775 (3.58%)	37	64,914 (5.42%)
		夕刊	1,415 (2.85%)	19	26,413 (2.20%)
	毎日新聞	朝刊	1,775 (3.58%)	30	53,757 (4.49%)
		夕刊	1,415 (2.85%)	11	15,565 (1.30%)
	読売新聞	朝刊	1,775 (3.58%)	37	65,929 (5.50%)
		夕刊	1,415 (2.85%)	19	27,357 (2.28%)
	日本経済新聞	朝刊	1,775 (3.58%)	39	69,986 (5.84%)
		夕刊	1,415 (2.85%)	17	24,527 (2.05%)
	産経新聞	朝刊	1,775 (3.58%)	32	56,800 (4.74%)
		夕刊	1,415 (2.85%)	15	21,225 (1.77%)
ブロック	北海道新聞	朝刊	1,775 (3.58%)	36	63,900 (5.33%)
		夕刊	1,415 (2.85%)	12	16,980 (1.42%)
	中日新聞	朝刊	1,775 (3.58%)	36	63,900 (5.33%)
		夕刊	1,415 (2.85%)	16	22,640 (1.89%)
	西日本新聞	朝刊	1,775 (3.58%)	36	63,900 (5.33%)
		夕刊	1,415 (2.85%)	12	16,980 (1.42%)
地方	河北新報	朝刊	1,775 (3.58%)	34	60,350 (5.04%)
		夕刊	1,415 (2.85%)	8	11,320 (0.94%)
	新潟日報	朝刊	1,775 (3.58%)	32	56,800 (4.74%)
		夕刊	1,415 (2.85%)	8	11,320 (0.94%)
	京都新聞	朝刊	1,775 (3.58%)	29	51,954 (4.34%)
		夕刊	1,415 (2.85%)	12	16,386 (1.37%)
	神戸新聞	朝刊	1,775 (3.58%)	30	53,836 (4.49%)
		夕刊	1,415 (2.85%)	13	18,041 (1.51%)
	中国新聞	朝刊	1,775 (3.58%)	32	56,800 (4.74%)
		夕刊	1,415 (2.85%)	8	11,320 (0.94%)
	高知新聞	朝刊	1,775 (3.58%)	26	46,150 (3.85%)
		夕刊	1,415 (2.85%)	8	11,320 (0.94%)
	愛媛新聞	朝刊	1,775 (3.58%)	28	49,700 (4.15%)
		夕刊	1,775 (3.58%)	32	56,800 (4.74%)
	琉球新報	朝刊	1,415 (2.85%)	8	11,320 (0.94%)

表 4.5: 書籍の NDC・判型ごとの冊数と累計

⇒ 参照元 2.2.1 節

	判型	冊数	冊数累計	
0. 総記	21cm	464	464	20.6%
	19cm	464	928	41.3%
	20cm	301	1,229	54.7%
	24cm	270	1,499	66.7%
	26cm	179	1,678	74.6%
	22cm	115	1,793	79.8%
	30cm	115	1,908	84.9%
	23cm	98	2,006	89.2%
	18cm	79	2,085	92.7%
	15cm	71	2,156	95.9%
	その他	92	2,248	100.0%
1. 哲学	19cm	1,273	1,273	35.20%
	20cm	711	1,984	54.87%
	22cm	424	2,408	66.59%
	21cm	422	2,830	78.26%
	15cm	263	3,093	85.54%
	18cm	223	3,316	91.70%
	16cm	100	3,416	94.47%
	26cm	76	3,492	96.57%
	23cm	31	3,523	97.43%
	27cm	26	3,549	98.15%
	その他	67	3,616	100.00%
2. 歴史	21cm	1,067	1,067	20.8%
	20cm	890	1,957	38.2%
	19cm	852	2,809	54.9%
	22cm	681	3,490	68.2%
	26cm	614	4,104	80.1%
	15cm	183	4,287	83.7%
	18cm	181	4,468	87.2%
	27cm	140	4,608	90.0%
	16cm	137	4,745	92.7%
	30cm	112	4,857	94.8%
	その他	264	5,121	100.0%
3. 社会科学	21cm	3,606	3,606	27.1%
	19cm	2,295	5,901	44.4%
	22cm	2,148	8,049	60.6%
	20cm	1,559	9,608	72.3%
	26cm	1,321	10,929	82.2%
	30cm	976	11,905	89.6%
	18cm	416	12,321	92.7%
	15cm	323	12,644	95.1%
	16cm	160	12,804	96.3%
	27cm	104	12,908	97.1%
	その他	384	13,292	100.0%
4. 自然科学	26cm	1,498	1,498	24.8%
	21cm	1,341	2,839	47.0%
	19cm	1,088	3,927	65.0%
	20cm	366	4,293	71.1%
	22cm	365	4,658	77.1%
	30cm	362	5,020	83.1%
	18cm	269	5,289	87.6%
	27cm	201	5,490	90.9%
	28cm	112	5,602	92.7%
	15cm	77	5,679	94.0%
	その他	362	6,041	100.0%
5. 技術・工学	21cm	1,430	1,430	22.26%
	26cm	1,150	2,580	40.16%
	30cm	999	3,579	55.71%
	19cm	645	4,224	65.75%
	22cm	462	4,686	72.95%
	24cm	346	5,032	78.33%
	20cm	261	5,293	82.39%
	27cm	196	5,489	85.45%
	18cm	150	5,639	87.78%
	15cm	130	5,769	89.80%
	その他	655	6,424	100.00%
6. 産業	21cm	744	744	23.60%
	19cm	608	1,352	42.89%
	30cm	548	1,900	60.28%
	22cm	287	2,187	69.38%
	26cm	284	2,471	78.39%
	20cm	236	2,707	85.88%
	18cm	99	2,806	89.02%
	27cm	58	2,864	90.86%
	15cm	57	2,921	92.67%
	16cm	45	2,966	94.10%
	その他	186	3,152	100.00%
7. 芸術・美術	21cm	1,419	1,419	26.97%
	19cm	929	2,348	44.62%
	26cm	595	2,943	55.93%
	20cm	507	3,450	65.56%
	22cm	365	3,815	72.50%
	30cm	334	4,149	78.85%
	18cm	228	4,377	83.18%
	15cm	170	4,547	86.41%
	16cm	141	4,688	89.09%
	31cm	95	4,783	90.90%
	その他	479	5,262	100.00%
8. 言語	21cm	232	232	24.63%
	19cm	231	463	49.15%
	18cm	110	573	60.83%
	22cm	105	678	71.97%
	26cm	60	738	78.34%
	20cm	59	797	84.61%
	15cm	44	841	89.28%
	16cm	31	872	92.57%
	17cm	19	891	94.59%
	27cm	19	910	96.60%
	その他	32	942	100.00%
9. 文学	20cm	3,665	3,665	24.12%
	19cm	2,816	6,481	42.66%
	15cm	2,686	9,167	60.34%
	16cm	1,666	10,833	71.31%
	22cm	1,405	12,238	80.56%
	18cm	1,205	13,443	88.49%
	21cm	747	14,190	93.40%
	17cm	607	14,797	97.40%
	26cm	92	14,889	98.01%
	23cm	65	14,954	98.43%
	その他	238	15,192	100.00%

表 4.6: 雑誌の分野・判型ごとの冊数と累計

⇒ 参照元 2.2.1 節

	判型	冊数	冊数累計	
1.総合	A4変	342	342	57.0%
	AB	94	436	72.7%
	B5	77	513	85.5%
	A4	31	544	90.7%
	A5	27	571	95.2%
	B5変	6	577	96.2%
	B6	5	582	97.0%
	A6	3	585	97.5%
	B4変	3	588	98.0%
	B6変	3	591	98.5%
	その他	9	600	100.0%
2.教育・学芸	A5	56	56	43.4%
	B5	56	112	86.8%
	A4変	9	121	93.8%
	AB	5	126	97.7%
	A4	2	128	99.2%
	A5変	1	129	100.0%
3.政治・経済・商業	A4変	30	30	71.4%
	B5	8	38	90.5%
	A5	2	40	95.2%
	B6	2	42	100.0%
4.産業	A4変	4	4	44.4%
	B5	3	7	77.8%
	菊倍	1	8	88.9%
	A4	1	9	100.0%
5.工業	A4変	90	90	82.6%
	B5	7	97	89.0%
	A4	7	104	95.4%
	AB	3	107	98.2%
	AB変	1	108	99.1%
	A5	1	109	100.0%
6.厚生・医療	A4変	8	8	40.0%
	B5	7	15	75.0%
	A4	2	17	85.0%
	AB	2	19	95.0%
	A5	1	20	100.0%

表 4.7: 目視計測と OCR 計測の結果の誤差 — 書籍

⇒ 参照元 2.2.3 節

	書籍タイトル	調査ページ	目視計測 文字数	OCR計測 文字数	誤差
1	世界遺産・大自然と秘境のミステリー	116	381	395	103.67%
		223	426	429	100.70%
2	もっと知りたい坂本竜馬	152	791	794	100.38%
		191	743	742	99.87%
3	自衛隊はどのようにして生まれたか	176	620	620	100.00%
		65	687	690	100.44%
4	交通事故 被害者の損害賠償 〔5訂版〕	147	493	494	100.20%
		247	547	551	100.73%
5	民事再生QA500	87	1276	1283	100.55%
		181	1391	1399	100.58%
6	あなたにもできる会社設立の本	38	317	319	100.63%
		208	539	541	100.37%
7	Q&A 非営利法人の＜会計・税務＞ 要点解説	287	957	959	100.21%
		195	642	645	100.47%
8	びっくり、ふしぎ 写真で科学 3 動物の目、人間の目	7	765	889	116.21%
		9	648	835	128.86%
9	アークザラット 精霊の黄昏 ザ・コ ンプリートガイド	37	991	991	100.00%
		43	700	716	102.29%
10	二十世紀名句手帖 1 人事篇	123	180	179	99.44%
		222	561	575	102.50%
11	ばあちゃん まるくなる	88	508	510	100.39%
		115	575	578	100.52%
12	古事記講義	207	907	910	100.33%
		185	676	679	100.44%
13	鉄道の怖い話	183	512	515	100.59%
		22	352	353	100.28%
14	現代中国少数民族詩集	141	468	469	100.21%
		87	430	431	100.23%
15	英語世界へのアプローチ	335	1489	1490	100.07%
		196	1091	1150	105.41%

表 4.8: 目視計測と OCR 計測の結果の誤差 — 雑誌

⇒ 参照元 2.2.3 節

	雑誌タイトル	調査ページ	目視計測 文字数	OCR計測 文字数	誤差
1	Animage	24	932	939	100.75%
		19	181	188	103.87%
2	HOBBY JAPAN	157	760	819	107.76%
		134	3338	3159	94.64%
3	My Birthday	117	2997	2451	81.78%
		90	1480	1478	99.86%
4	Newton	92	494	517	104.66%
		71	1758	1824	103.75%
5	Option	169	2397	2404	100.29%
		102	1897	1830	96.47%
6	蛍雪時代	92	888	889	100.11%
		102	1531	1530	99.93%
7	月刊テレビジョン(首都圏版)	85	4105	4136	100.76%
		10	2013	2039	101.29%
8	壮快	172	1253	1129	90.10%
		262	869	828	95.28%
9	日経PC21	99	1877	1818	96.86%
		161	1019	1022	100.29%
10	日経TRENDY	150	1682	1715	101.96%
		175	1687	1717	101.78%

表 4.9: 書籍の文字数計測結果 (1/5)

⇒ 参照元 3.1.1 節

NDC-判型	書籍タイトル	出版社	頁1	文字数	頁2	文字数	頁3	文字数	頁4	文字数	頁5	文字数	総ページ数	キャラクタ密度
1 0-18	情報検索のスキル未知の問題をどう解くか	中央公論新社	36	518	71	550	107	591	143	409	178	586	214	531
2 0-19	勝ち残るSEの絶対条件	ナツメ社	39	13	77	642	116	537	154	684	193	432	231	462
3 0-19	へんな本あります	風塵社	37	502	74	691	111	488	147	825	184	764	221	654
4 0-20	古書発見女たちの本を追って	影書房	42	216	85	218	127	640	169	251	212	624	254	390
5 0-21	情報処理教科書 情報セキュリティアドミニストレータ午前 平成15年度	翔泳社	62	667	294	794	249	555	142	591	129	553	320	632
6 0-21	写真著作権	草の根出版会	36	559	71	495	107	363	143	613	178	1419	214	690
7 0-22	丸山眞男集 4	岩波書店	50	0	383	210	8	637	15	827	245	837	392	502
8 0-23	3DCGアニメーション基礎から最先端まで	技術評論社	47	541	93	758	140	586	186	466	233	605	279	591
9 0-24	仕事に使える Access2002 活用ガイド 基礎編	毎日コミュニケーションズ	73	451	46	427	72	633	88	605	347	222	360	468
10 0-26	図書館利用教育ハンドブック 大学図書館版	日本図書館協会	35	1334	70	1304	105	1229	139	1158	174	107	209	1026
11 0-30	図書館における自己点検・評価等のあり方に関する調査研究報告書	日本図書館協会	24	631	48	405	72	432	95	841	119	538	143	569
12 1-15	聖書時代史	岩波書店	42	539	84	543	127	500	169	558	211	580	253	544
13 1-16	ポケット判風水 2004大開運術	廣済堂出版	53	249	28	364	237	500	18	441	244	168	350	344
14 1-16	鉄道の怖い話	廣済堂出版	183	512	166	434	111	421	22	352	120	423	247	428
15 1-16	恋愛論アンソロジー	中央公論新社	264	493	99	428	357	548	56	492	353	453	440	483
16 1-18	親鸞と暗闇をやぶる力宗教という生きる知恵	講談社	36	479	73	600	109	577	145	576	182	479	218	542
17 1-19	魂の目覚め	IBC21	36	511	71	655	107	465	142	427	178	532	213	518
18 1-19	心理学	岩波書店	35	455	69	510	104	338	139	731	173	733	208	553
19 1-19	日本人の情感はどこからくるのか	草思社	35	210	70	492	105	472	140	468	175	572	210	443
20 1-19	神の発明	講談社	35	620	69	681	104	592	139	20	173	456	208	474
21 1-20	ドロシーおばさんの心をひらこう	扶桑社	41	39	82	479	123	488	164	0	205	499	246	301
22 1-20	ドゥルーズへのまなざし	筑摩書房	41	914	82	630	123	857	163	588	204	802	245	758
23 1-21	あなたのところを科学する3[改訂3版]	北大路書房	122	863	62	999	132	0	217	633	164	930	272	685
24 1-21	ビレン色彩学の謎を解く人類の福祉向上のための色彩心理学と色彩療法	青娥書房	37	826	74	799	112	830	149	775	186	784	223	803
25 1-22	清沢満之全集 3 哲学論集	岩波書店	120	726	122	460	142	435	369	517	277	264	378	480
26 1-22	民数記	河出書房新社	85	1067	150	818	13	723	113	832	94	716	250	831
27 2-15	石原莞爾その虚飾	講談社	54	449	107	561	161	642	215	641	268	287	322	516
28 2-16	石原慎太郎の値打ち。	宝島社	48	501	95	541	143	458	190	245	238	425	285	434
29 2-18	真説・川中島合戦封印された戦国最大の白兵戦	洋泉社	38	619	76	0	114	639	151	555	189	520	227	467
30 2-19	もっと知りたい坂本龍馬	日本実業出版社	152	791	69	629	191	743	26	833	116	788	264	757
31 2-19	大江戸開府四百年事情	講談社	40	778	79	618	119	30	158	0	198	95	237	304
32 2-19	旅ちから	幻冬舎	40	306	79	290	119	432	158	412	198	334	237	355
33 2-20	死の骨董	以文社	6	653	27	763	28	712	130	712	10	590	224	686
34 2-20	世界悪女大全	文藝春秋	207	591	108	538	233	606	19	507	133	14	352	451
35 2-20	聖徳太子の真実	平凡社	13	769	378	791	53	705	401	744	83	677	440	737
36 2-20	中島知久平	日本経済評論社	116	591	2	476	246	220	179	9	223	673	252	394
37 2-20	物語スウェーデン史	新評論	114	611	22	324	63	660	157	547	61	614	240	551
38 2-20	老いは生のさなかにあり	幻冬舎	9	21	114	524	184	527	2	171	96	554	254	359
39 2-21	トラベルストーリー 7 釜山・慶州・済州島(ブサン・キョンジュ・チェジュドー)	昭文社	36	506	101	273	125	104	173	1187	185	1609	232	736
40 2-21	鳥めぐりフェリーで行こう! スロー・トラベル	東京書籍	35	748	69	160	104	733	138	712	173	138	207	498
41 2-21	江戸大名データファイル 127人	新人物往来社	35	765	69	759	104	743	139	669	173	688	208	725
42 2-22	大津事件考	大阪大学出版会	116	380	118	604	143	17	39	888	94	614	164	501
43 2-22	同志社山脈	晃洋書房	45	1029	161	961	5	896	103	1021	33	1061	268	994
44 2-26	環境考古学マニュアル	同成社	162	1105	109	1645	65	1400	50	1200	310	1392	400	1348
45 2-26	歩いて巡る 東京	実業之日本社	84	1224	81	181	15	1270	25	934	118	1995	128	1121
46 2-27	日光・鎌倉・箱根	学習研究社	15	1149	29	547	44	824	58	714	73	1024	87	852
47 3-15	いまだからズバリ! この資格	河出書房新社	46	514	147	494	85	538	10	481	41	42	224	414
48 3-18	企業再生の基礎知識	岩波書店	90	534	167	486	104	538	1	71	8	409	184	408
49 3-18	韓国人から見た北朝鮮独裁国家のルーツ	PHP研究所	36	534	72	449	109	555	145	474	181	510	217	504
50 3-19	Q&A過労死・過労自殺110番[全訂増補版]	民事法研究会	214	664	168	436	238	112	158	479	157	469	252	432

表 4.10: 書籍の文字数計測結果 (2/5)

⇒ 参照元 3.1.1 節

NDC-判型	書籍タイトル	出版社	頁1	文字数	頁2	文字数	頁3	文字数	頁4	文字数	頁5	文字数	総ページ数	キャラクタ密度
51 3-19	あなたにもできる会社設立の本	フォレスト出版	38	317	208	539	64	316	60	499	206	522	224	439
52 3-19	お笑い日中戦闘宣言!	実業之日本社	78	477	143	525	181	423	122	517	129	537	226	496
53 3-19	クイア・スタディーズ	岩波書店	7	734	10	729	102	711	65	590	108	748	128	702
54 3-19	クマのプーさんと学ぶマネジメント	ダイヤモンド社	254	0	67	544	6	398	164	524	8	255	256	344
55 3-19	ヤクザ・リセッション さらに失われる10年	光文社	11	610	133	691	202	709	228	649	85	623	280	656
56 3-19	見えない目で生きるとのこと	明石書店	24	563	103	420	29	587	142	422	118	577	200	514
57 3-19	司法書士試験なるほどわかった民法のABC	東京リーガルマインド	4	119	335	337	63	228	209	229	233	362	374	255
58 3-19	実用六法 平成16年版	成美堂出版	574	1581	165	2508	682	1426	86	1600	402	1089	912	1641
59 3-19	新・教育原理〔改訂版〕	有斐閣	71	483	67	688	151	711	249	190	190	660	274	546
60 3-20	いのちのひみつ	KTC中央出版	249	749	112	606	200	373	267	688	196	692	280	622
61 3-20	シークレット・パワー	リベルタ出版	239	791	203	857	299	594	62	1028	119	724	480	799
62 3-20	マルクスと『資本論』2	新日本出版社	337	735	93	647	285	628	107	814	313	743	384	713
63 3-20	家族革命前夜(イブ)	集英社インターナショナル(集英社)	15	480	104	436	185	581	222	0	146	594	240	418
64 3-20	自衛隊はどのようにして生まれたか	学研	176	620	214	0	190	644	8	292	206	600	304	431
65 3-20	新現代商法入門〔第2版〕	法律文化社	275	688	118	684	7	481	256	721	259	632	332	641
66 3-20	選挙制度と政党	信山社出版	224	697	252	493	98	743	73	545	96	504	290	596
67 3-21	Q&A 非営利法人の(会計・税務)要点解説	清文社	287	957	404	448	59	533	195	642	255	315	516	579
68 3-21	すぐに役立つ夫婦・親子の法律と手続き	三修社	52	596	190	381	51	608	142	231	78	519	224	467
69 3-21	会計諸則集〔最新増補4版〕	税務経理協会	914	602	106	825	367	33	319	742	389	951	1012	631
70 3-21	建設技術者のためのPFI入門	技術書院	2	0	155	592	4	384	177	861	168	711	184	510
71 3-21	実務 契約法講義	民法法研究会	404	882	143	880	458	875	119	666	370	795	499	820
72 3-21	商業登記申請Memo〔補訂版〕	新日本法規出版	301	544	168	639	116	609	142	583	135	707	422	616
73 3-21	地方分権	山川出版社	121	13	44	855	194	882	31	12	234	1015	200	555
74 3-21	非公開企業の企業価値評価	税務経理協会	94	917	144	391	140	634	262	172	35	582	280	539
75 3-21	シャヒード、100の命パレスチナで生きて死ぬこと	「シャヒード、100の命」展実行委員会	36	1100	72	1806	108	1315	143	21	179	10	215	850
76 3-22	5年後こうなる	PHPソフトウェア・グループ(PHP研究所)	44	406	174	539	128	503	167	441	68	529	256	484
77 3-22	アマルフィ海法研究試論	関西大学出版部	34	1019	272	1328	279	1044	207	1039	201	963	320	1079
78 3-22	グローバル企業の管理会計	中央経済社	254	914	236	805	152	1032	150	1022	98	867	292	928
79 3-22	バリュー・アット・リスクのすべて〔新版〕	シグマベイスキャピタル	367	518	259	822	215	890	4	644	166	787	626	732
80 3-22	会社経営と取締役の責任	成文堂	33	644	148	589	86	826	117	873	169	760	230	738
81 3-22	企業会計とディスクロージャー〔第2版〕	東京大学出版会	6	899	193	21	125	97	118	822	102	539	312	476
82 3-22	経済変動理論	中央経済社	151	74	81	549	45	316	8	507	164	792	210	448
83 3-22	現代会社法講義	成文堂	150	799	403	283	220	973	329	1011	55	1095	608	832
84 3-22	国際比較調査の方法と解析	慶應義塾大学出版会	1	12	155	750	45	36	139	793	350	773	520	473
85 3-22	自由の倫理学	勁草書房	383	1317	129	769	173	893	178	913	10	685	424	915
86 3-22	商法総論総則	有斐閣	193	887	254	1097	220	1163	298	1154	174	1115	348	1083
87 3-22	消費者法〔第2版〕	有斐閣	163	1006	248	1070	229	1189	6	1012	196	848	448	1025
88 3-22	新商法下の会社経営と決算・監査	税務経理協会	19	66	157	940	122	576	144	341	152	661	272	517
89 3-22	人権主体としての個と集団	日本評論社	155	909	102	778	76	747	86	792	186	807	344	807
90 3-22	税務六法 法令編 平成15年版	ぎょうせい	331	2456	2669	2923	1860	2914	2747	2585	2785	2884	3504	2752
91 3-26	医師・医療機関のための保険診療ルールBook 2003年版	医学通信社	79	1711	235	416	142	1538	119	547	12	743	240	991
92 3-26	対話で学ぶ行政法	有斐閣	145	1641	43	1394	290	1949	246	1808	99	1882	318	1735
93 3-26	法学	ミネルヴァ書房	10	1066	119	1187	77	1311	150	1016	163	838	272	1084
94 3-26	民事再生QA500	信山社出版	87	1276	181	1391	323	1101	123	1275	73	1361	374	1281
95 3-30	新しいアーバンライフ	日本経済新聞社	26	1714	17	1311	168	470	54	1551	75	1718	170	1353
96 3-30	変貌する都市圏	日経産業消費研究所	27	2823	54	2423	82	3430	109	795	136	3258	163	2546
97 3-30	共用品白書	ぎょうせい	47	1343	93	1325	140	1119	186	942	233	890	279	1124
98 4-17	骨格バランスDIEI	祥伝社	31	161	47	247	14	374	23	332	81	340	104	291
99 4-18	ふしぎの博物誌動物・植物・地学の32話	中央公論新社	36	236	72	588	108	611	144	230	180	620	216	457

表 4.11: 書籍の文字数計測結果 (3/5)

⇒ 参照元 3.1.1 節

NDC-判型	書籍タイトル	出版社	頁1	文字数	頁2	文字数	頁3	文字数	頁4	文字数	頁5	文字数	総ページ数	キャラクタ密度
100	4-19 うつ病者からの手紙	新曜社	247	530	138	642	31	625	36	596	298	0	320	479
101	4-19 かたちのパノラマ	丸善	79	438	155	424	179	365	161	315	7	389	192	386
102	4-19 遺伝子は語る	河出書房新社	90	594	134	619	80	532	45	625	54	343	176	543
103	4-20 心はなぜ腰痛を選ぶのか サノー博士の心身症治療プログラム	春秋社	44	586	88	664	133	613	177	593	221	702	265	632
104	4-21 ひっくり、ふしぎ 写真で科学 3 動物の目、人間の目	大月書店	7	765	6	58	31	750	27	252	9	648	36	495
105	4-21 肝臓・胆のう・膵臓病の食事療法〔第3版〕	医歯薬出版	29	1200	192	1210	75	73	198	191	207	1403	238	815
106	4-21 認定病院評価結果の情報提供 2003 近畿版〔新版〕	日本医療機能評価機構	611	251	683	17	576	998	1113	960	709	1078	1234	661
107	4-21 アスペルガー症候群と高機能自閉症その基本的理解のために	エンパワメント研究所	30	698	61	804	91	858	121	722	152	679	182	752
108	4-22 看護師・看護学生のためのレビューブック〔第5版〕	メディックメディア	882	386	285	434	82	389	1049	1948	997	546	1072	741
109	4-22 工学基礎 離散数学とその応用	数理工学社(サイエンス社)	193	927	99	957	95	647	32	923	144	854	224	862
110	4-26 医学・保健学のためのやさしい統計学	金原出版	42	216	119	236	30	187	140	515	182	5718	208	1374
111	4-26 栄養の生化学1-2-3〔改訂第3版〕	南江堂	57	1411	183	521	66	883	131	523	45	535	230	775
112	4-26 生活習慣病の理解〔第2版〕	文光堂	145	1604	71	1486	83	1443	139	1368	84	1576	264	1495
113	4-26 食品機能学	建帛社	27	702	54	1257	82	915	109	670	136	1427	163	994
114	4-27 栄養学・食品学を学ぶための食品化学実験	地人書館	22	534	45	110	67	589	89	870	112	265	134	474
115	4-30 細胞と組織の地図帳	講談社	26	786	53	795	79	423	105	89	132	809	158	580
116	5-15 世界遺産・大自然と秘境のミステリー	PHP研究所	116	381	148	454	94	495	223	426	71	396	290	430
117	5-16 毎日楽ちんナチュラル家事	光文社	38	468	76	100	114	516	152	600	190	296	228	396
118	5-18 「伝統の町並み」の歩き方 日本各地に残る国宝級の町61	青春出版社	35	206	71	213	106	584	141	235	177	466	212	341
119	5-19 電池がわかる本	オーム社	201	100	207	484	62	766	168	721	151	56	216	425
120	5-19 日本プラモデル興亡史わたしの模型人生	文春ネスコ	35	89	69	583	104	565	138	539	173	196	207	394
121	5-20 「青色」に挑んだ男たち	日本経済新聞社	3	386	33	566	2	7	120	593	121	622	284	435
122	5-20 岩波講座 インターネット 4 ネットワークアプリケーション	岩波書店	122	716	45	1024	16	0	134	711	66	765	240	643
123	5-21 アマチュア局用 電波法令抄録	CQ出版	177	1379	44	1078	59	722	120	710	185	1402	192	1058
124	5-21 おなかの中から始める子育て	サンマーク出版	155	459	110	384	18	402	15	283	39	415	160	389
125	5-21 評伝・桑沢洋子	桑沢学園(そしえて)	90	573	34	509	185	598	186	627	130	554	272	572
126	5-21 トコトンやさしいパーソナルロボットの本	日刊工業新聞社	27	318	53	763	80	893	106	894	133	493	159	672
127	5-22 岩波講座 環境経済・政策学 8 環境の評価とマネジメント	岩波書店	145	968	222	1168	70	891	142	929	30	721	272	935
128	5-22 新編 建築材料	市ヶ谷出版社	181	664	137	647	113	753	152	729	127	685	240	696
129	5-24 アマチュア無線 自作電子回路	CQ出版	111	845	117	1187	150	846	83	1256	141	170	200	861
130	5-26 血液サラサラ! 発芽玄米もっとおいしいレシピ100	主婦と生活社	5	749	7	781	35	633	17	798	69	914	96	775
131	5-26 佐藤貴美枝のニットソーイング	雄鶏社	2	155	11	185	16	383	4	185	8	154	40	212
132	5-26 詳解Webサービス構築	ソフトバンクパブリッシング	304	896	366	1127	55	977	118	1214	129	1385	400	1120
133	5-27 新しいPC橋の設計	山海堂	25	215	50	412	76	555	101	796	126	746	151	545
134	5-29 ディスク丸ごと完全コピー	祥伝社	34	574	2	833	43	464	12	718	38	592	96	636
135	5-30 建築行政法規と秩序を学ぶ	学芸出版社	23	1140	45	1492	68	2003	90	2167	113	2231	135	1807
136	5-30 成功した1000万円台の家	ニューハウス出版	27	898	54	1103	81	82	107	135	134	825	161	609
137	5-30 民家再生の魅力全国・事例選集	相模書房	28	596	56	313	85	696	113	155	141	601	169	472
138	6-18 交通事故紛争	文藝春秋	202	501	235	598	25	591	242	491	46	613	256	559
139	6-19 交通事故 被害者の損害賠償〔5訂版〕	新星出版社	147	493	65	389	247	547	74	743	156	347	288	504
140	6-19 黒ダイズ	農山漁村文化協会	31	445	10	443	61	568	192	321	56	473	250	450
141	6-19 資産活用の切り札! ハワイ別荘・不動産投資のすすめ	郁朋社	66	532	206	449	199	567	162	684	107	707	240	588

表 4.12: 書籍の文字数計測結果 (4/5)

⇒ 参照元 3.1.1 節

NDC-判型	書籍タイトル	出版社	頁1	文字数	頁2	文字数	頁3	文字数	頁4	文字数	頁5	文字数	総ページ数	キャラクタ密度
142 6-20	ネットでビジネスを成功させる方法	税務経理協会	199	491	190	604	292	648	278	449	147	366	320	512
143 6-21	マーケティング・データ解析 Excel/Accessによる	朝倉書店	30	425	61	226	91	420	121	171	152	558	182	360
144 6-21	「地域一番店」はここが違う	経林書房	30	429	60	381	91	611	121	649	151	613	181	537
145 6-22	ブランディング360° 思考	東洋経済新報社	178	882	93	508	47	529	143	854	133	539	240	662
146 6-22	国土交通六法 国土編 平成16年版	新日本法規出版	2964	2286	740	1700	763	1705	2839	2213	2115	1399	3128	1861
147 6-22	内航海運ハンドブック 2003年版	成山堂書店	30	983	268	1296	249	1164	173	1402	20	431	280	1055
148 6-26	輸出統計品目表 2004年版	日本関税協会	15	1548	75	473	240	1106	128	1604	306	1542	730	1255
149 6-27	イモとヒト人類の生存を支えた根栽農耕	平凡社	59	1187	119	239	178	595	237	1155	297	2440	356	1123
150 6-30	地球のすばらしい樹木たち 巨樹・奇樹・神木	早川書房	32	1853	64	0	96	0	127	0	159	0	191	371
151 6-30	絵葉書に見る懐かしの鉄道	ほおずき書籍	30	431	60	820	91	369	121	594	151	348	181	512
152 7-15	歌右衛門の疎開	岩波書店	45	507	91	586	136	499	181	459	227	467	272	504
153 7-16	ちひろ美術館ものがたり	講談社	76	544	24	391	120	448	73	543	152	538	256	493
154 7-18	廃墟の美学	集英社	36	652	71	398	107	72	142	467	178	667	213	451
155 7-19	ドラゴンクエストモンスターズ キャラバンハート	集英社	106	431	211	1075	35	728	205	810	133	795	230	768
156 7-19	すらすら読める南方録	講談社	37	464	74	431	111	408	148	513	185	381	222	439
157 7-19	クラブ文化が人を育てる学校・地域を再生するスポーツクラブ論	大修館書店	37	606	74	656	112	644	149	564	186	633	223	621
158 7-19	すらすら読める風姿花伝	講談社	37	572	74	518	111	362	148	362	185	514	222	466
159 7-20	指揮台の神々	音楽之友社	319	721	355	822	259	775	37	813	143	849	472	796
160 7-20	心が変われば	朝日新聞社	216	1071	19	414	210	1016	189	357	39	387	224	649
161 7-20	戦後史のなかの映画	スペース伽耶(星雲社)	23	719	399	715	447	753	10	731	285	725	488	729
162 7-20	禅と合気道(オンデマンド版)	人文書院	10	337	86	479	119	563	54	507	113	335	176	444
163 7-21	アークザラッド 精霊の黄昏 ゼン・コンプリートガイド(角川書店)	メディアワークス(角川書店)	37	991	36	801	244	1276	270	781	43	700	272	910
164 7-21	ザンボット3・ダイターン3大全	双葉社	58	1047	13	45	59	1230	95	1078	185	1525	223	985
165 7-21	子どものスポーツ医学入門 現場の疑問にきちんと答える スポーツと「発育・栄養・障害・クスリ」に関するQ&A75	山海堂	29	464	58	643	88	700	117	32	146	643	175	496
166 7-21	英語でおりがみを楽しむ	日本文芸社	29	33	58	396	88	238	117	329	146	389	175	277
167 7-22	ゴシック美術サン・ドニからの旅立ち	八坂書房	42	560	84	750	127	604	169	832	211	448	253	639
168 7-26	気軽に楽しむ絵手紙	西東社	48	102	58	139	8	191	35	115	95	147	152	139
169 7-26	茶道具の箱と箱書	淡交社	19	24	37	166	56	921	74	1028	93	352	111	498
170 7-29	極上カメラ倶楽部 OLYMPUS E-1	双葉社	102	640	98	1307	21	1892	74	824	82	116	112	956
171 7-30	今竹七郎とその時代グラフィックデザイン、モダン絵画の先駆者 1905-2000	誠文堂新光社	21	999	41	2115	62	2460	83	1665	103	315	124	1511
172 7-31	かたちのなかの源氏物語	弓立社	33	631	65	1265	98	621	130	1348	163	29	195	779
173 8-15	第2言語習得のメカニズム	筑摩書房	44	542	88	654	132	662	175	546	219	599	263	601
174 8-16	通勤電車でメキメキ上達! 大人の漢字力!	宝島社	37	128	74	370	111	416	147	461	184	350	221	345
175 8-18	「ことば」の課外授業ハダシの学者の言語学1週間	洋泉社	36	507	72	521	109	538	145	468	181	577	217	522
176 8-19	フランス語の常識日常表現は文化の鏡	白水社	37	344	73	661	110	856	146	722	183	577	219	632
177 8-20	アメリカを知るための英語、アメリカから離れるための英語	文藝春秋	115	33	242	727	53	579	104	766	8	703	288	562
178 8-21	日本語運用文法	凡人社	178	779	155	1021	146	651	170	1015	195	1086	237	910
179 8-22	英語世界へのアプローチ カリスマ受験講師出口汪の日本語トレーニング	小学館	161	802	335	1489	196	1091	172	922	195	956	352	1052
180 8-26	あやかし 下	双葉社	29	385	58	306	88	729	117	511	146	602	175	507
181 9-15	カンピュセス王の秘宝	角川書店	82	419	49	603	63	203	297	629	22	619	336	495
182 9-15	ジャンキー	河出書房新社	171	533	26	0	70	421	62	616	177	411	296	396
183 9-15	ハートに火をつけて	集英社	28	337	78	347	199	445	39	150	88	455	224	347
184 9-15	ママと少年 いけない童貞いじり	マドンナ社(二見書房)	167	403	28	457	82	552	163	380	17	534	224	465
185 9-15	怪奇探偵小説名作選 10	筑摩書房	276	623	373	457	267	536	229	652	196	661	528	586
186 9-15	新・はっぴいセブン 4 秋の夜はハードボイルド	集英社	149	0	120	310	145	353	244	52	108	378	260	219

表 4.13: 書籍の文字数計測結果 (5/5)

⇒ 参照元 3.1.1 節

NDC-判型	書籍タイトル	出版社	頁1	文字数	頁2	文字数	頁3	文字数	頁4	文字数	頁5	文字数	総ページ数	キャラクタ密度
188 9-15	地にはマのつく星が降る!	角川書店	163	480	90	553	157	552	161	404	80	599	224	518
189 9-15	冥海の霊剣 仙姫幻想	講談社	224	574	220	442	20	468	152	488	223	275	288	449
190 9-15	容赦なく恋に落ちなさい	角川書店	6	413	199	502	30	459	110	495	204	496	226	473
191 9-16	黄泉津比良坂、血祭りの館	徳間書店	304	378	358	404	218	414	407	564	356	520	512	456
192 9-16	秘密の顔を持つ女	扶桑社	59	616	119	629	178	554	237	495	297	643	356	587
193 9-16	カリブ海の秘密	早川書房	59	425	118	465	177	473	236	520	295	498	354	476
194 9-16	探偵家族	早川書房	60	494	119	376	179	555	239	160	298	363	358	390
195 9-16	ライフ・オブ・デビッド・ゲイル アラシ・パーカー作品	新潮社	59	569	117	585	176	406	235	543	293	530	352	527
196 9-18	「サラ川」傑作選 さんびょうし	講談社	21	17	110	111	170	207	97	103	182	191	200	126
197 9-18	てのひらを、たいように 上	パラダイム	196	532	137	603	128	601	61	473	225	517	244	545
198 9-18	幻神伝	光文社	26	691	441	641	317	374	50	668	478	618	544	598
199 9-18	深夜管理人・魅惑の秘貝探訪	パラダイム	87	652	65	643	41	574	140	598	35	625	208	618
200 9-18	六本木夜遊民(ナイトバード)	白泉社	191	457	56	392	111	616	165	622	22	545	204	526
201 9-19	ベア・ナックル	講談社	240	453	171	462	185	559	297	482	328	540	360	499
202 9-19	駅に住人	文藝書房	53	242	7	437	18	471	32	442	100	233	183	365
203 9-19	現代中国少数民族詩集	土曜美術社出版販売	21	293	6	0	141	468	87	430	90	431	168	324
204 9-19	三太郎の六月	碧天舎	134	466	145	284	71	499	12	372	132	356	300	395
205 9-19	透明人間	偕成社	173	582	56	662	267	607	39	354	145	284	291	498
206 9-19	二十世紀名句手帖 1 人事篇	河出書房新社	123	180	222	561	202	170	95	189	42	173	240	255
207 9-19	甦る男	早川書房	82	779	307	738	474	798	7	3	159	641	520	592
208 9-19	女子高生になれなかった 少年ある性同一性障害者の 青春時代	青弓社	41	662	82	619	123	592	164	811	205	588	246	654
209 9-20	おばかさんに乾杯	小峰書店	190	382	187	512	163	437	164	464	204	461	235	451
210 9-20	ばあちゃん まるくなる	溪水社	88	508	1	2	115	575	92	515	158	371	230	394
211 9-20	維新物語	文藝書房	129	498	134	399	72	532	204	515	26	447	220	478
212 9-20	決定版 三島由紀夫全集 32 私の遍歴時代 他	新潮社	245	613	548	748	263	183	352	767	267	602	728	583
213 9-20	骨ん中	小学館	248	725	212	590	1	3	170	667	43	362	352	469
214 9-20	志賀直哉、上高畑の「サロ ン」をめぐる考察	創元社	73	460	191	624	144	593	91	587	125	606	238	574
215 9-20	斜光の中で	文芸社	31	348	108	460	47	455	38	342	232	335	272	388
216 9-20	水の精霊 第II部 赤光	ポプラ社	366	462	39	493	273	573	468	208	462	526	471	452
217 9-20	世界のかなたの森	晶文社	35	651	263	596	172	602	242	633	208	619	276	620
218 9-20	銭とり橋	幻冬舎	239	484	193	548	276	574	119	396	41	495	284	499
219 9-20	本格的	原書房	41	693	144	513	197	725	160	657	216	678	288	653
220 9-20	養安先生、呼ばれ!	恒文社21(恒文社)	129	544	364	0	401	563	220	577	473	587	486	454
221 9-21	吉原幸子	思潮社	35	1395	69	530	104	732	139	1077	173	1021	208	951
222 9-21	私の武蔵野探勝吟行入門	日本放送出版協会	35	832	69	421	104	623	138	546	173	356	207	556
223 9-22	アリアドネの糸	英宝社	284	849	173	786	201	798	449	458	63	783	500	735
224 9-22	ときめきのまほう王国へ!!	ポプラ社	120	206	73	0	56	176	72	238	15	285	126	181
225 9-22	ナポコフ伝(ロシア時代)2	みすず書房	130	1071	27	988	339	1553	319	2066	169	1090	376	1354
226 9-22	ふたごもり	砂子屋書房	21	61	28	25	96	59	240	55	50	52	268	50
227 9-22	古事記講義	文藝春秋	18	0	207	907	216	836	194	615	185	676	280	607

表 4.14: 書籍の文字数調査のまとめ — 発行年・NDC・判型別の冊数

⇒ 参照元 3.1.1 節

2001年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		53	51	1	83	406	350	461	109	158
1. 哲学		275	152	26	190	1,043	754	323	518	35
2. 歴史		247	150	5	210	781	928	918	728	30
3. 社会科学		335	192	10	412	2,191	1,537	3,393	2,079	44
4. 自然科学		109	80	10	254	974	366	1,160	422	20
5. 技術工学		113	98	13	179	684	343	1,421	445	110
6. 産業		61	44	12	127	589	249	729	310	9
7. 芸術		150	149	12	188	810	528	1,373	303	57
8. 言語		47	31	17	148	355	75	410	114	9
9. 文学		2,551	1,641	525	1,006	2,130	3,665	686	1,050	88
n. 記録なし		107	23	19	167	253	164	426	431	46

2002年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		65	67	7	50	359	290	499	124	162
1. 哲学		264	126	28	220	1,150	775	357	519	29
2. 歴史		219	138	8	223	720	887	1,033	670	28
3. 社会科学		329	178	19	460	2,158	1,546	3,545	2,143	60
4. 自然科学		110	66	15	291	991	392	1,235	362	35
5. 技術工学		115	98	18	165	642	322	1,448	418	140
6. 産業		52	51	18	109	580	250	737	256	13
7. 芸術		157	153	7	176	909	601	1,384	292	64
8. 言語		49	25	13	136	338	60	309	123	3
9. 文学		2,624	1,678	588	1,142	2,361	3,768	648	1,154	78
n. 記録なし		86	12	41	151	232	168	508	322	65

2003年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		80	46	4	80	467	303	466	115	98
1. 哲学		268	101	23	223	1,275	711	424	426	31
2. 歴史		190	139	7	183	859	893	1,069	684	30
3. 社会科学		331	168	23	417	2,301	1,561	3,629	2,151	63
4. 自然科学		90	74	14	269	1,092	377	1,350	374	30
5. 技術工学		142	107	23	194	656	303	1,445	465	90
6. 産業		63	47	15	102	615	244	750	293	7
7. 芸術		197	146	9	241	940	519	1,435	371	54
8. 言語		44	31	19	110	231	59	232	106	4
9. 文学		2,738	1,682	607	1,212	2,842	3,670	755	1,409	66
n. 記録なし		12	12	16	30	61	11	465	15	26

2004年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		78	34	9	72	497	275	511	118	80
1. 哲学		305	112	31	214	1,429	750	454	471	50
2. 歴史		218	125	8	185	904	845	1,042	725	24
3. 社会科学		212	160	14	468	2,326	1,337	3,650	2,046	46
4. 自然科学		114	71	20	268	1,087	349	1,384	376	35
5. 技術工学		193	88	18	152	587	282	1,408	427	86
6. 産業		57	38	5	120	649	216	782	308	14
7. 芸術		187	140	18	232	904	509	1,402	322	64
8. 言語		65	24	13	84	212	48	260	122	13
9. 文学		3,080	1,767	625	1,265	3,148	3,671	762	1,445	64
n. 記録なし		10	5	16	58	90	10	455	11	51

2005年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		52	46	8	80	524	192	398	104	61
1. 哲学		293	124	43	207	1,205	573	385	354	29
2. 歴史		190	116	12	148	777	721	815	549	36
3. 社会科学		208	133	12	419	2,087	1,095	3,128	1,809	30
4. 自然科学		99	57	18	244	884	305	1,205	319	96
5. 技術工学		155	75	20	131	539	242	1,272	390	77
6. 産業		44	31	6	79	604	167	705	234	15
7. 芸術		163	119	16	214	766	394	1,265	256	53
8. 言語		36	19	17	93	215	45	200	111	5
9. 文学		2,989	1,528	614	1,033	3,248	3,342	642	1,190	60
n. 記録なし		9	7	7	69	96	6	299	7	32

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
411	2	239	9	56	5	106	3	0	2,503
2	2	98	29	2	2	21	2	0	3,474
34	30	711	147	11	41	138	97	0	5,206
30	6	1,262	118	43	9	1,058	94	0	12,813
25	23	1,307	193	69	33	382	42	0	5,469
425	69	1,396	168	230	133	1,274	110	0	7,211
28	6	367	53	28	67	547	52	0	3,278
89	22	859	112	105	126	314	100	1	5,298
5	0	70	6	6	1	23	0	0	1,317
10	3	87	26	1	2	16	1	0	13,488
26	14	445	231	110	231	173	48	0	2,914

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
384	5	205	6	33	6	82	2	0	2,346
6	4	74	26	0	2	28	27	0	3,635
29	11	631	123	10	17	104	98	1	4,950
27	9	1,211	87	33	14	1,020	98	1	12,938
33	15	1,389	187	80	35	417	80	0	5,733
423	73	1,202	155	116	61	1,209	87	0	6,692
31	7	293	51	19	33	534	49	0	3,083
72	32	676	97	53	71	332	123	0	5,199
1	0	64	2	0	0	20	0	0	1,143
13	12	86	21	1	1	28	14	0	14,217
159	11	934	198	419	422	621	68	0	4,417

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
270	3	179	13	3	1	115	5	0	2,248
5	3	76	26	2	3	15	4	0	3,616
19	10	619	140	13	38	112	116	0	5,121
48	17	1,322	104	30	58	976	92	1	13,292
56	22	1,498	202	112	43	362	76	0	6,041
348	70	1,153	197	63	71	1,000	97	0	6,424
27	4	284	58	11	40	548	44	0	3,152
85	30	597	92	48	60	335	103	0	5,262
1	3	60	19	3	2	16	2	0	942
21	21	92	46	5	1	16	9	0	15,192
181	4	1,328	25	589	547	1,103	3	0	4,428

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
242	7	162	20	3	5	71	1	0	2,185
6	3	91	32	2	6	23	14	0	3,993
13	12	576	154	12	33	155	106	0	5,137
48	25	1,338	103	26	55	1,002	68	0	12,924
59	14	1,537	195	121	49	458	75	0	6,212
297	76	1,146	174	49	61	906	61	1	6,012
40	4	286	79	18	20	499	46	0	3,181
77	24	635	113	60	53	343	117	0	5,200
5	9	60	23	1	2	31	2	0	974
19	13	95	34	4	1	13	5	0	16,011
221	6	1,427	26	642	631	1,234	4	0	4,897

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
149	8	143	7	7	3	66	2	0	1,850
8	1	81	10	3	5	23	3	2	3,349
8	3	446	122	9	29	139	90	0	4,210
64	11	1,084	107	15	33	737	46	1	11,019
51	15	1,345	143	97	57	303	52	0	5,290
207	52	920	176	53	52	614	63	0	5,038
31	7	245	48	11	29	350	32	0	2,638
78	26	524	93	45	70	265	81	0	4,428
8	1	43	9	1	9	23	0	0	835
24	7	72	33	6	2	16	2	0	14,808
153	11	1,108	25	436	574	1,042	3	0	3,884

表 4.15: 書籍の文字数調査のまとめ — 発行年・NDC・判型別のページ数

⇒ 参照元 3.1.1 節

2001年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		13,210	14,459	126	17,907	88,358	87,468	114,816	35,626	62,226
1. 哲学		67,808	43,339	3,761	40,518	232,361	197,275	72,122	196,123	13,773
2. 歴史		74,439	47,816	588	52,300	174,636	264,789	200,866	308,899	14,360
3. 社会科学		91,578	57,266	1,915	90,252	519,786	417,993	793,236	711,574	41,530
4. 自然科学		27,506	22,009	1,676	52,051	210,933	92,836	245,160	133,638	5,847
5. 技術工学		28,882	24,158	2,199	33,649	145,921	82,088	313,079	131,605	34,070
6. 産業		15,485	11,338	2,359	24,873	128,317	63,144	159,255	91,282	3,130
7. 芸術		38,484	41,213	1,708	39,160	175,704	136,428	269,645	88,851	15,163
8. 言語		12,695	10,428	3,138	30,375	77,737	21,023	86,363	40,094	3,482
9. 文学		743,805	555,732	99,784	242,881	449,275	951,953	151,970	304,017	31,630
n. 記録なし		19,030	2,360	2,787	35,725	42,031	26,293	63,659	68,890	4,959

2002年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		14,244	18,559	1,267	10,329	74,263	71,528	119,526	40,058	62,351
1. 哲学		63,766	31,114	4,516	47,569	250,823	201,486	82,824	194,030	12,852
2. 歴史		67,691	43,966	1,601	48,020	165,552	242,846	221,339	278,284	10,657
3. 社会科学		87,534	51,542	3,656	98,951	512,554	411,430	841,482	737,560	54,150
4. 自然科学		26,902	18,309	2,364	60,293	218,040	97,559	259,382	112,306	11,817
5. 技術工学		28,809	25,903	2,408	31,299	135,983	72,712	320,041	116,964	41,221
6. 産業		12,508	13,061	3,116	22,816	126,255	64,000	158,537	74,965	3,875
7. 芸術		40,180	41,886	725	35,013	199,591	155,979	284,287	81,854	10,404
8. 言語		11,960	7,500	2,174	29,846	77,761	15,525	62,689	39,107	875
9. 文学		784,524	579,716	111,222	272,011	487,249	971,822	123,932	318,941	31,115
n. 記録なし		13,115	1,995	10,846	28,730	42,474	30,487	79,116	54,901	12,014

2003年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		16,707	13,777	514	16,991	102,140	70,976	120,492	36,427	37,333
1. 哲学		62,960	25,249	3,157	47,567	277,776	183,400	92,025	157,187	13,854
2. 歴史		56,451	42,127	1,161	42,513	189,291	241,282	238,395	273,430	11,663
3. 社会科学		91,942	48,971	4,280	91,813	543,466	418,146	846,597	742,770	56,789
4. 自然科学		21,065	20,447	2,194	54,993	236,486	96,030	293,920	115,376	9,746
5. 技術工学		36,842	26,807	4,411	34,972	139,172	67,977	313,545	129,470	24,571
6. 産業		14,738	11,023	2,716	20,941	134,811	59,924	162,627	86,418	2,030
7. 芸術		47,816	41,212	2,025	48,339	201,529	134,524	283,848	103,029	11,794
8. 言語		11,941	8,480	4,562	21,529	49,550	14,992	48,056	34,136	877
9. 文学		796,650	593,481	118,516	288,525	589,521	943,202	152,276	373,353	19,503
n. 記録なし		1,747	2,780	2,757	4,681	10,740	1,239	79,788	1,768	8,059

2004年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		16,542	10,100	1,863	15,354	98,159	60,490	127,381	35,106	30,391
1. 哲学		71,828	29,115	6,024	45,530	311,665	192,472	104,447	183,341	27,544
2. 歴史		58,535	39,745	1,321	44,236	193,625	236,321	234,199	291,891	9,853
3. 社会科学		58,126	45,355	2,490	100,943	543,226	361,612	858,596	720,467	38,911
4. 自然科学		26,716	19,521	3,695	56,060	234,829	90,577	301,386	115,988	10,650
5. 技術工学		52,310	23,343	3,641	28,267	122,887	62,565	295,040	120,518	25,779
6. 産業		14,121	10,102	761	24,773	140,292	50,377	168,705	90,338	3,182
7. 芸術		45,398	38,532	3,234	49,031	194,478	129,723	277,445	88,569	11,214
8. 言語		15,530	5,432	3,373	17,695	46,750	12,461	54,699	46,141	2,120
9. 文学		894,829	613,482	124,267	300,934	610,977	932,870	143,559	363,457	20,989
n. 記録なし		1,708	1,192	2,921	9,215	14,439	1,345	84,996	1,496	15,332

2005年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		11,003	12,883	1,755	16,170	102,764	41,178	95,761	31,903	19,390
1. 哲学		66,502	29,670	6,168	43,736	258,147	150,189	83,820	128,021	12,256
2. 歴史		55,944	37,263	1,669	33,532	174,265	209,806	187,765	212,851	15,533
3. 社会科学		59,788	38,344	2,613	91,687	486,766	298,607	738,818	631,921	20,987
4. 自然科学		18,819	15,374	3,875	48,625	185,177	77,595	258,522	103,615	11,592
5. 技術工学		40,604	18,703	3,589	23,898	112,830	55,819	259,394	109,336	18,651
6. 産業		10,786	7,692	902	15,488	126,749	41,464	146,646	74,577	2,410
7. 芸術		39,810	27,614	2,135	44,351	163,619	105,537	247,617	68,255	8,396
8. 言語		8,518	5,192	3,682	19,768	46,588	10,812	44,124	38,664	2,560
9. 文学		876,651	542,543	124,568	249,561	602,130	824,029	116,533	301,286	17,217
n. 記録なし		1,644	1,529	1,269	10,117	19,383	946	60,230	1,015	8,434

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
146,374	378	54,983	2,926	9,742	1,108	17,073	935	0	667,715
158	222	17,412	10,990	367	807	2,726	246	0	900,008
4,483	2,035	113,794	74,592	1,638	6,978	23,384	26,209	0	1,391,806
9,231	902	274,091	61,515	7,845	1,247	162,542	37,134	0	3,279,637
5,536	6,164	313,513	74,057	17,338	7,019	65,494	14,651	0	1,295,428
122,333	8,319	233,839	61,331	35,449	20,817	220,437	41,240	0	1,539,416
5,886	1,506	63,841	22,059	4,844	13,731	82,372	16,034	0	709,456
22,241	2,467	129,987	24,068	14,231	18,217	40,352	25,309	84	1,083,312
1,147	0	12,549	2,042	767	141	2,618	0	0	304,599
2,413	492	12,988	9,182	159	240	2,123	246	0	3,558,890
2,887	1,127	54,342	15,528	14,145	23,383	20,169	3,405	0	400,720

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
152,378	1,068	48,382	1,078	4,758	1,634	11,867	594	0	633,884
1,241	2,694	11,348	9,559		395	4,779	10,329	0	929,325
3,106	2,079	98,074	54,858	1,161	3,265	16,239	28,389	270	1,287,397
6,543	2,455	254,802	40,978	6,948	5,199	161,126	36,520	162	3,313,592
7,123	2,645	344,309	69,932	21,611	6,328	71,719	26,820	0	1,357,459
121,446	7,288	204,961	49,789	19,589	9,400	209,678	26,925	0	1,424,416
7,030	1,261	50,081	18,998	4,287	6,622	77,966	14,215	0	659,593
18,820	4,934	100,002	20,402	7,196	10,356	41,991	16,193	0	1,069,813
255	0	9,918	782	0	0	2,450	0	0	260,842
1,898	969	12,040	5,546	223	226	5,779	3,632	0	3,710,845
45,225	1,092	126,985	15,739	59,851	54,714	84,804	4,655	0	666,743

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
99,606	597	42,628	3,617	464	207	21,339	491	0	584,306
907	289	11,279	8,491	164	295	1,943	837	0	887,380
4,275	1,151	95,161	61,376	1,942	3,233	17,830	29,563	0	1,310,844
11,486	4,058	284,643	40,831	4,686	7,011	161,151	36,563	187	3,395,390
11,034	4,847	369,920	70,954	35,035	11,927	68,104	23,927	0	1,446,005
129,333	8,214	192,553	64,435	11,563	10,053	182,483	30,452	0	1,406,853
5,651	1,462	47,353	17,527	2,066	8,588	81,239	15,379	0	674,493
18,109	3,267	83,002	17,951	5,655	8,711	40,183	16,476	0	1,067,470
255	174	12,640	4,425	562	318	1,653	668	0	214,818
2,110	1,991	15,026	7,236	692	48	2,007	1,365	0	3,905,502
53,465	437	178,466	3,355	81,513	76,043	143,998	349	0	651,185

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
86,058	1,843	35,132	8,412	603	1,192	11,980	263	0	540,869
1,385	430	16,424	8,908	1,023	490	3,180	3,571	0	1,007,377
2,776	1,346	89,759	76,332	2,142	3,090	23,301	29,616	0	1,338,088
11,591	4,892	280,123	35,026	4,830	5,848	180,201	23,723	0	3,275,960
11,676	3,233	361,241	71,108	28,215	6,126	85,119	25,818	0	1,451,958
77,291	10,570	187,488	59,340	8,224	7,077	163,207	14,885	678	1,263,110
7,350	1,129	49,819	30,284	3,078	3,187	79,340	13,830	0	690,668
17,067	3,039	92,124	20,838	8,065	7,470	43,191	16,012	0	1,045,430
907	1,417	9,063	6,498	64	207	3,022	292	0	225,671
2,323	1,131	15,826	5,993	510	68	2,328	1,187	0	4,034,730
70,225	1,166	191,039	3,055	87,402	85,205	164,053	426	0	735,215

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
57,975	1,845	27,160	3,186	1,260	309	8,160	317	0	433,019
2,105	127	14,334	5,213	763	235	2,778	388	787	805,239
1,529	323	65,202	67,386	1,258	4,122	24,900	27,689	0	1,121,037
15,814	3,539	227,204	31,805	2,104	4,858	122,800	16,833	49	2,794,537
11,723	2,214	320,131	50,414	30,623	7,272	58,972	16,565	0	1,221,108
54,455	5,471	152,910	51,620	10,354	5,575	105,257	19,074	0	1,047,540
5,969	1,270	41,292	15,956	1,810	6,730	56,927	7,435	0	564,103
17,154	4,409	77,370	21,187	6,701	8,182	31,741	13,428	0	887,506
920	325	6,130	731	127	424	2,345	0	0	190,910
3,145	1,135	10,400	5,978	639	580	1,698	218	0	3,678,311
45,547	1,875	146,349	2,912	59,263	76,036	132,956	487	0	569,992

表 4.16: 書籍の文字数調査のまとめ — 発行年・NDC・判型別の推計総文字数

⇒ 参照元 3.1.1 節

2001年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		5,426,074	6,391,023	59,631	9,505,036	49,286,092	34,095,026	75,881,894	17,891,377	36,788,011
1. 哲学		36,887,552	18,138,816	1,779,950	21,968,860	115,483,417	104,476,840	53,651,556	128,617,463	9,101,130
2. 歴史		38,410,524	20,752,144	278,280	24,403,180	82,416,550	140,285,212	131,152,107	230,778,443	9,489,016
3. 社会科学		37,894,976	25,312,145	906,302	41,154,912	313,191,856	252,037,836	490,678,162	630,378,663	27,442,816
4. 自然科学		11,298,227	9,728,198	487,381	23,787,307	98,955,701	58,635,218	166,892,670	107,057,402	3,863,668
5. 技術工学		12,430,813	9,566,568	1,040,710	11,467,579	59,813,018	44,245,432	210,639,551	107,323,878	22,513,286
6. 産業		6,360,541	5,011,509	1,116,432	13,899,032	65,937,829	32,304,470	71,394,017	108,875,084	2,068,288
7. 芸術		19,380,542	20,309,766	808,337	17,668,992	100,739,888	89,285,305	179,866,697	56,758,019	10,019,635
8. 言語		7,624,617	3,597,660	1,485,106	15,861,825	49,129,784	11,806,517	78,624,875	42,178,888	2,300,888
9. 文学		323,703,936	270,730,401	47,224,275	117,272,662	201,196,577	477,325,100	114,479,001	177,947,230	20,900,946
n. 記録なし		7,816,668	1,043,144	1,318,990	18,023,977	22,519,159	14,908,920	38,086,225	43,369,011	3,276,882

2002年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		5,850,794	8,203,264	599,627	5,482,633	41,423,901	27,881,614	78,994,733	20,117,128	36,861,911
1. 哲学		34,688,704	13,022,246	2,137,265	25,791,912	124,659,031	106,706,986	61,612,774	127,244,874	8,492,537
2. 歴史		34,928,556	19,081,244	757,697	22,406,132	78,129,507	128,659,811	144,519,611	207,905,976	7,042,092
3. 社会科学		36,221,569	22,782,079	1,730,257	45,121,656	308,834,287	248,080,535	520,522,066	653,399,487	35,782,049
4. 自然科学		11,050,131	8,092,761	687,451	27,553,901	102,289,832	61,618,264	176,574,297	89,968,337	7,808,615
5. 技術工学		12,399,394	10,257,588	1,139,622	10,666,699	55,739,432	39,191,768	215,323,585	95,384,142	27,238,631
6. 産業		5,137,724	5,773,093	1,474,694	12,749,581	64,878,236	32,742,400	71,072,137	89,413,254	2,560,581
7. 芸術		20,234,648	20,641,421	343,117	15,797,866	114,435,500	102,080,457	189,633,643	52,288,335	6,874,911
8. 言語		7,183,176	2,587,500	1,028,878	15,585,581	49,144,952	8,718,840	57,072,066	41,140,564	578,196
9. 文学		341,424,845	282,414,447	52,637,480	131,337,791	218,202,283	487,287,748	93,357,976	186,682,546	20,560,636
n. 記録なし		5,387,052	881,810	5,133,032	14,494,860	22,756,507	17,287,044	47,333,916	34,562,376	7,938,791

2003年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		6,862,484	6,089,572	243,258	9,018,823	56,973,692	27,666,445	79,633,163	18,293,639	22,071,270
1. 哲学		34,250,240	10,567,548	1,494,098	25,790,827	138,054,672	97,128,640	68,457,398	103,083,235	9,154,654
2. 歴史		29,128,716	18,283,118	549,461	19,836,566	89,332,733	127,831,204	155,656,042	204,279,553	7,706,852
3. 社会科学		38,045,600	21,645,672	2,025,574	41,866,728	327,460,004	252,130,091	523,686,091	658,014,991	37,525,887
4. 自然科学		8,652,554	9,037,778	638,015	25,131,801	110,943,465	60,652,548	200,086,040	92,427,714	6,440,108
5. 技術工学		15,856,797	10,615,572	2,087,572	11,918,458	57,046,603	36,639,603	210,953,076	105,582,785	16,236,394
6. 産業		6,053,707	4,872,276	1,285,388	11,701,831	69,274,879	30,657,118	72,905,684	103,073,629	1,341,414
7. 芸術		24,080,138	20,309,274	958,362	21,810,557	115,546,652	88,039,232	189,340,808	65,814,925	7,793,416
8. 言語		7,171,765	2,925,600	2,159,035	11,242,444	31,315,600	8,419,507	43,750,182	35,911,072	579,517
9. 文学		346,702,080	289,120,204	56,089,475	139,311,411	264,002,242	472,937,203	114,709,511	218,530,978	12,887,485
n. 記録なし		717,589	1,228,788	1,304,792	2,361,658	5,754,224	702,550	47,735,964	1,113,027	5,325,347

2004年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		6,794,709	4,464,301	881,693	8,149,903	54,753,090	23,579,002	84,186,103	17,630,233	17,967,159
1. 哲学		39,074,432	12,185,598	2,850,948	24,686,366	154,897,505	101,933,171	77,698,123	120,235,028	18,200,937
2. 歴史		30,204,060	17,249,330	625,183	20,640,518	91,378,092	125,202,866	152,916,334	218,071,766	6,510,813
3. 社会科学		24,052,539	20,047,364	1,178,430	46,030,008	327,315,394	218,041,704	531,108,406	638,256,912	25,712,194
4. 自然科学		10,973,731	8,628,477	1,074,506	25,619,420	110,166,112	57,208,433	205,168,520	92,917,987	7,037,467
5. 技術工学		22,514,224	9,243,828	1,723,158	9,633,394	50,371,381	33,722,535	198,502,912	98,282,429	17,034,634
6. 産業		5,800,271	4,465,185	360,155	13,843,152	72,091,382	25,772,873	75,630,452	107,749,144	2,102,650
7. 芸術		22,862,433	18,988,570	1,530,539	22,122,787	111,503,961	84,897,217	185,069,687	56,577,877	7,410,155
8. 言語		9,327,318	1,874,040	1,596,323	9,240,329	29,546,000	6,998,098	49,797,970	48,540,332	1,400,885
9. 文学		389,429,581	298,863,891	58,811,222	145,302,973	273,610,775	467,756,566	108,142,995	212,738,651	13,869,426
n. 記録なし		701,570	526,876	1,382,407	4,649,152	7,736,055	762,655	50,851,832	941,792	10,131,309

2005年		15cm以下	16cm	17cm	18cm	19cm	20cm	21cm	22cm	23cm
0. 総記		4,519,537	5,694,415	830,580	8,583,036	57,321,759	16,051,184	63,288,445	16,021,687	11,463,368
1. 哲学		36,177,088	12,417,884	2,919,099	23,713,659	128,299,059	79,540,094	62,353,698	83,956,172	8,098,704
2. 歴史		28,867,104	16,172,142	789,879	15,646,031	82,241,462	111,155,219	122,598,027	159,020,982	10,264,129
3. 社会科学		24,740,274	16,948,431	1,236,641	41,809,272	293,295,986	180,051,489	457,016,397	559,814,601	13,868,105
4. 自然科学		7,729,998	6,795,462	1,126,850	22,221,625	86,872,703	49,009,002	175,988,852	83,005,977	7,659,936
5. 技術工学		17,475,962	7,406,388	1,698,548	8,144,438	46,249,017	30,086,441	174,520,283	89,163,508	12,324,488
6. 産業		4,430,403	3,399,941	426,885	8,654,694	65,132,086	21,212,982	65,741,402	88,950,474	8,950,516
7. 芸術		20,048,316	13,608,179	1,010,421	20,011,171	93,810,954	69,068,690	165,172,920	43,601,294	5,548,035
8. 言語		5,115,911	1,791,240	1,742,562	10,322,850	29,443,616	6,072,019	40,170,490	40,674,528	1,691,635
9. 文学		381,518,515	264,305,248	58,953,675	120,498,033	269,648,867	413,181,874	87,784,309	176,348,722	11,376,908
n. 記録なし		675,281	675,833	600,573	5,104,229	10,384,927	536,410	36,034,706	638,983	5,573,145

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
68,444,482	273,409	56,434,551	2,299,295	7,959,896	939,944	9,721,366	851,631	0	382,248,739
109,344	160,574	13,138,399	8,636,107	299,865	684,598	2,397,735	224,065	0	515,756,270
3,102,460	1,471,926	140,490,072	63,522,547	1,338,361	5,919,612	20,568,099	23,872,075	0	938,250,607
6,388,314	652,421	348,794,502	48,339,410	6,409,914	1,057,861	272,116,980	33,822,947	0	2,536,580,019
3,831,189	4,458,452	363,565,350	35,073,395	14,166,360	5,954,393	38,012,718	13,344,644	0	959,112,273
105,304,246	6,017,174	164,248,514	33,413,129	28,964,314	13,243,775	212,163,265	37,562,835	0	1,079,958,087
4,073,406	1,089,297	80,094,919	24,776,669	3,957,887	11,648,351	36,367,238	14,604,328	0	483,579,299
15,391,884	1,784,393	41,400,860	18,912,995	11,627,723	17,411,809	60,963,802	19,710,649	57,799	682,099,095
793,781	0	6,357,323	1,604,634	626,693	119,614	2,302,740	0	0	224,414,946
1,669,917	355,866	9,800,225	7,215,353	129,914	203,598	1,867,348	224,065	0	1,772,246,415
1,997,948	815,165	41,004,300	12,202,135	11,557,455	19,836,383	17,740,249	3,101,393	0	258,618,003

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
71,251,953	772,490	49,659,285	847,109	3,887,619	1,386,163	6,757,070	541,036	0	360,518,330
858,834	1,948,584	8,562,747	7,511,606	0	335,088	4,203,513	9,408,015	0	537,184,714
2,149,507	1,503,751	121,082,160	46,717,073	948,618	2,769,781	14,283,500	25,857,695	189,156	858,931,868
4,528,083	1,775,714	324,248,285	32,201,127	5,677,002	4,410,442	269,746,407	33,263,694	122,738	2,548,447,478
4,929,472	1,913,142	399,277,932	33,119,795	17,657,700	5,368,201	41,625,708	24,428,595	0	1,013,964,132
104,540,717	5,271,447	143,964,606	27,125,047	16,005,584	5,980,280	201,808,086	24,524,232	0	996,560,860
4,865,112	912,088	62,831,623	21,338,554	3,502,779	5,617,608	34,421,989	12,947,520	0	432,238,970
13,024,381	3,568,787	31,850,637	16,032,198	5,879,636	9,898,265	63,440,003	12,611,108	0	678,634,912
176,473	0	5,024,459	614,507	0	0	2,154,971	0	0	191,010,162
1,313,511	700,883	9,084,902	4,358,130	182,207	191,721	5,083,093	3,308,153	0	1,838,128,351
31,297,961	789,849	95,817,802	12,367,942	48,902,457	46,415,254	74,591,902	4,239,937	0	470,198,491

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
46,575,766	431,813	43,753,379	2,842,293	379,120	175,603	12,150,427	447,220	0	333,607,966
627,689	209,035	8,510,682	6,672,355	133,999	250,256	1,709,024	762,369	0	506,856,721
2,958,514	832,524	117,485,771	52,267,802	1,586,750	2,742,635	15,682,911	26,927,015	0	873,088,165
7,948,886	2,935,172	362,222,450	32,085,612	3,828,790	5,947,607	269,788,261	33,302,860	141,679	2,620,601,954
7,636,080	3,505,859	428,977,728	33,603,814	28,626,047	10,117,972	39,527,562	21,793,549	0	1,087,798,636
111,329,846	5,941,227	135,249,227	35,104,188	9,447,780	6,395,719	175,633,805	27,736,747	0	973,775,399
3,910,775	1,057,472	59,409,074	19,686,326	1,688,067	7,285,415	35,867,019	14,007,731	0	444,077,805
12,532,333	2,363,037	26,436,137	14,106,165	4,620,531	8,325,974	60,708,476	12,831,509	0	675,617,526
176,473	125,855	6,403,424	3,477,231	459,193	269,767	1,453,946	608,438	0	156,449,050
1,460,226	1,440,100	11,338,019	5,686,157	565,412	40,720	1,765,317	1,243,290	0	1,937,829,829
37,000,453	316,084	134,663,305	2,636,409	66,601,827	64,509,178	126,657,761	317,881	0	498,946,836

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
40,240,721	1,333,051	36,059,485	6,610,276	492,693	1,011,203	6,821,412	239,550	0	311,214,584
958,489	311,021	12,392,893	7,000,040	835,863	415,679	2,797,064	3,252,592	0	579,725,751
1,921,131	973,569	110,816,461	65,004,331	1,750,164	2,621,324	20,495,094	26,975,289	0	893,356,324
8,021,552	3,538,408	356,470,524	27,523,956	3,946,448	4,961,005	301,680,501	21,607,739	0	2,559,493,082
8,080,376	2,338,445	418,913,126	33,676,749	23,053,630	5,196,839	49,403,068	23,515,938	0	1,082,972,822
66,532,093	7,645,334	131,691,571	32,328,432	6,719,584	4,502,387	157,081,297	13,557,779	434,572	861,521,544
5,086,568	816,611	62,502,917	34,014,989	2,514,941	2,703,612	35,028,610	12,596,848	0	463,080,360
11,811,217	2,198,124	29,341,494	16,374,813	6,589,670	7,139,826	65,252,963	12,470,146	0	662,141,479
627,689	1,024,923	4,591,316	5,106,226	52,292	175,603	2,658,091	265,964	0	172,823,399
1,607,632	818,058	11,941,667	4,709,389	416,706	57,686	2,047,662	1,081,161	0	1,991,206,040
48,599,211	843,374	144,150,388	2,400,665	71,413,552	72,281,532	144,297,738	388,016	0	562,058,122

24cm	25cm	26cm	27cm	28cm	29cm	30cm	31cm以上	null	計
27,109,110	1,334,498	27,877,024	2,503,607	1,029,508	262,132	4,646,304	288,735	0	248,824,929
1,456,765	91,860	10,815,863	4,096,454	623,424	199,356	2,443,473	353,404	531,301	458,087,357
1,058,144	233,628	80,498,389	57,385,918	1,027,874	3,496,796	21,901,542	25,220,110	0	737,577,377
10,944,079	2,559,776	289,128,450	24,992,846	1,719,115	4,121,163	205,583,573	15,332,086	37,125	2,143,199,410
8,112,902	1,601,397	371,239,914	23,876,070	25,021,135	6,169,019	34,227,349	15,087,982	0	925,746,173
46,874,864	3,957,202	107,403,984	28,122,576	8,459,943	3,546,815	101,306,354	17,373,267	0	704,114,077
4,130,846	918,597	51,804,943	17,921,779	1,478,897	5,709,227	25,133,271	6,772,058	0	373,411,003
11,871,426	3,189,052	24,642,345	16,649,062	5,475,186	7,820,356	47,954,303	10,457,726	0	559,939,435
636,686	235,074	3,105,458	574,431	103,768	359,690	2,062,615	0	0	144,102,572
2,176,497	820,951	7,847,424	4,697,602	522,108	492,029	1,493,527	198,562	0	1,801,864,850
31,520,801	1,356,197	110,429,101	2,288,293	48,422,019	64,503,240	116,945,438	443,577	0	436,132,755

表 4.17: 雑誌の文字数計測結果

⇒ 参照元 3.1.2 節

	分野	判型	タイトル	出版社名	採用巻号	計測対象 頁数	総かたろ ぐ頁数
1	1. 総合	A4	TVガイド	東京ニュース通信社	2003.4.4	149	152
2	1. 総合	A4	ボクシングマガジン	ベースボール・マガジン社	2003.4	148	124
3	1. 総合	A4変型	animage	徳間書店	2003.11	143	220
4	1. 総合	A4変型	BACKSTAGE PASS	シンコー・ミュージック	2003.8	161	196
5	1. 総合	A4変型	dancyu	プレジデント社	2003.8	133	176
6	1. 総合	A4変型	L magazine	京阪神エルマガジン社	2003.8	156	184
7	1. 総合	A4変型	MONEY japan	角川SSコミュニケーションズ	2003.12	117	176
8	1. 総合	A4変型	Myojo	集英社	2003.10	167	188
9	1. 総合	A4変型	SCREEN	近代映画社	2003.12	182	216
10	1. 総合	A4変型	新しい住まいの設計	扶桑社	2003.12	214	328
11	1. 総合	A4変型	一個人	ベストセラーズ	2003.11	136	140
12	1. 総合	A4変型	ひよこクラブ	ベネッセコーポレーション	2003.8	168	340
13	1. 総合	A4変型	MORE	集英社	2003.8	291	452
14	1. 総合	A4変型	GOLF DIGEST	ゴルフダイジェスト社	2003.8	167	300
15	1. 総合	A4変型	AERA (アエラ)	朝日新聞社東京本社	2003.3.10	71	80
16	1. 総合	A4変型	すてきな奥さん	主婦と生活社	2003.2	177	280
17	1. 総合	A4変型	やさい畑	家の光協会	2003.秋	112	168
18	1. 総合	A4変型	私の家づくり	主婦と生活社	2003.3月	143	200
19	1. 総合	A4変型	日経マネー	日経ホーム出版社	2003.10	131	200
20	1. 総合	A4変型	美しい部屋	主婦と生活社	2003.6	149	228
21	1. 総合	A4変型	婦人公論	中央公論新社	2003.1.22	159	230
22	1. 総合	A5	文藝春秋	文藝春秋	2003.8	394	464
23	1. 総合	AB	BE-PAL	小学館	2003.8	170	250
24	1. 総合	AB	My Birthday	実業之日本社	2003.11	143	222
25	1. 総合	AB	月刊ザテレビジョン8版	角川書店	2003.10	159	180
26	1. 総合	AB	YOUNG GUITAR	シンコー・ミュージック	2003.10	149	224
27	1. 総合	AB	週刊女性	主婦と生活社	2003.9.23	156	235
28	1. 総合	B4変型	家庭画報	世界文化社	2003.11	317	370
29	1. 総合	B5	HOBBY JAPAN	ホビージャパン	2003.9	219	376
30	1. 総合	B5	趣味の園芸	日本放送出版協会(NHK出版)	2003.8	132	136
31	1. 総合	B5	アサヒカメラ	朝日新聞社東京本社	2003.10	337	352
32	1. 総合	B5	相撲	ベースボール・マガジン社	2003.10	112	202
33	2. 教育・学芸	A4変型	Newton	ニュートンプレス	2003.11	129	144
34	2. 教育・学芸	A5	小説宝石	光文社	2003.11	464	440
35	2. 教育・学芸	A5	俳句	角川書店	2003.10	269	280
36	2. 教育・学芸	A5	文学界	文藝春秋	2003.11	312	352
37	2. 教育・学芸	B5	蜚雪時代	旺文社	2003.11	181	200
38	2. 教育・学芸	B5	季刊銀花	文化出版局	2003.冬	154	200
39	2. 教育・学芸	B5	仏教芸術	毎日新聞社出版局	2003.9	81	112
40	3. 政治・経済・商業	A4変型	日経TRENDY	日経ホーム出版社	2003.10	125	275
41	3. 政治・経済・商業	A4変型	週刊ダイヤモンド	ダイヤモンド社	2003.6.14	143	136
42	3. 政治・経済・商業	B5	財界	財界研究所	2003.10.21	124	150
43	4. 産業	B5	農耕と園藝	誠文堂新光社	2003.2	171	264
44	5. 工業	A4変型	Option	三栄書房	2003.11	198	496
45	5. 工業	A4変型	日経PC21	日経BP社	2003.11	148	220
46	5. 工業	A4変型	ASAHIパソコン	朝日新聞社東京本社	2003.9.15	94	180
47	5. 工業	A4変型	月刊アスキー	アスキー	2003.10	221	500
48	5. 工業	A4変型	日経アーキテクチャ	日経BP社	2003.9.1	125	150
49	5. 工業	A4変型	日経エコロジー	日経BP社	2003.8	106	140
50	5. 工業	A4変型	日経ソフトウエア	日経BP社	2003.9	131	200
51	6. 厚生・医療	A4変型	ゆほびか	マキノ出版	2003.1	158	166
52	6. 厚生・医療	A4変型	日経メディカル	日経BP社	2003.11	163	200
53	6. 厚生・医療	B5	壮快	マキノ出版	2003.12	209	328

	頁1	文字数	頁2	文字数	頁3	文字数	頁4	文字数	頁5	文字数	キャラクタ 密度
1	32	1,397	65	1,894	96	3,874	128	1,693	160	1,176	1,968
2	32	884	64	2,135	96	2,092	128	2,908	192	2,630	2,534
3	37	2,376	24	932	19	181	45	4,460	16	450	1,092
4	10	17	173	3,341	142	3,219	22	1,906	28	4	1,394
5	8	999	160	1,615	74	1,194	35	1,029	94	1,339	933
6	74	3,289	120	7,061	160	8,448	98	3,127	25	1,287	3,936
7	146	1,347	57	2,109	11	414	121	263	60	1,602	762
8	74	870	33	989	44	349	168	1,416	18	127	666
9	70	862	11	1,238	190	1,650	87	717	85	2,523	1,178
10	63	1,486	102	318	7	1,362	137	503	178	713	572
11	17	340	44	891	58	697	80	798	125	980	720
12	172	1,058	271	1,449	55	1,662	183	2,388	61	521	699
13	316	3,713	54	1,010	65	481	19	694	390	979	885
14	21	816	221	246	169	510	190	1,320	125	360	362
15	15	2,193	30	2,372	44	2,555	61	1,125	75	2,115	1,831
16	38	860	76	842	114	1,415	152	1,315	228	3,031	942
17	24	1,014	48	500	72	1,093	96	658	120	478	500
18	31	445	62	812	93	1,490	124	689	155	295	533
19	28	1,240	56	1,893	84	1,299	113	1,779	140	1,541	1,013
20	32	485	66	1,000	96	1,020	128	815	160	961	560
21	35	755	70	1,785	103	1,575	140	155	173	1,798	839
22	306	1,055	320	1,241	446	1,354	100	1,206	264	1,216	1,031
23	39	926	138	2,053	38	666	180	847	60	931	738
24	117	2,997	90	1,480	69	5,774	3	2,009	138	3,210	1,993
25	17	0	148	4,850	85	4,105	10	2,013	61	2,307	2,345
26	36	1,642	64	2,937	96	2,785	128	554	160	626	1,137
27	34	1,160	68	1,048	102	1,549	136	2,103	170	584	855
28	388	146	440	2,664	138	689	371	935	216	416	831
29	200	303	157	760	134	3,338	682	745	17	726	684
30	85	1,384	92	0	53	1,103	133	1,299	26	1,028	934
31	73	47	156	797	219	302	292	684	342	1,351	608
32	29	197	58	4,315	87	1,390	116	1,679	145	1,451	1,003
33	92	494	71	1,758	85	147	115	1,716	57	157	765
34	493	1,291	51	6	155	569	422	1,143	99	1,211	890
35	93	876	296	901	205	612	213	1,040	242	470	749
36	295	1,158	54	1,167	21	1,156	124	1,329	224	916	1,015
37	92	888	102	1,531	9	957	205	1,147	199	1,929	1,168
38	40	74	80	522	120	65	160	1,100	200	1,682	529
39	18	1,793	36	1,793	54	1,110	72	743	90	1,006	937
40	150	1,682	175	1,688	200	1,019	65	1,899	127	1,779	733
41	31	2,134	63	1,337	93	1,221	124	190	155	1,280	1,295
42	27	949	54	1,175	82	213	108	1,827	135	645	798
43	37	55	74	867	111	1,376	148	725	185	897	507
44	169	2,397	102	1,897	47	1,395	163	2,128	359	2,097	792
45	99	1,877	161	1,019	190	1,358	102	1,063	197	1,211	878
46	20	2,316	41	1,930	66	1,061	80	1,749	100	1,067	846
47	48	1,761	96	2,078	144	2,420	192	1,767	240	2,465	929
48	27	720	54	1,144	81	1,259	108	2,250	137	362	958
49	23	1,694	46	1,966	69	1,721	91	1,994	118	1,692	1,373
50	28	42	56	2,598	84	1,235	112	2,110	140	1,952	1,037
51	34	1,024	68	1,277	102	1,378	136	1,375	163	1,399	1,231
52	35	1,841	70	2,068	103	2,199	133	724	175	1,247	1,316
53	172	1,253	177	360	262	869	163	330	131	1,190	510

表 4.18: 雑誌の文字数調査のまとめ — 発行年・分野・判型別の冊数

⇒ 参照元 3.1.2 節

2001年		A4判	A4変 型判	A5判	A5変 型判	A6判	AB・A5 判	AB判	AB変 型判
	1.総合	269	4,107	278	3	15	0	1,393	12
	2.教育・学芸	6	66	544	12	0	0	62	0
	3.政治・経済・商業	0	393	40	0	0	0	12	0
	4.産業	64	37	0	0	0	0	0	0
	5.工業	86	1,378	24	0	0	0	96	6
	6.厚生・医療	24	96	2	0	0	0	24	0

2002年		A4判	A4変 型判	A5判	A5変 型判	A6判	AB・A5 判	AB判	AB変 型判
	1.総合	305	4,274	284	3	15	12	1,197	64
	2.教育・学芸	12	88	520	12	0	0	62	0
	3.政治・経済・商業	0	453	40	0	0	0	42	0
	4.産業	52	28	0	0	0	0	0	0
	5.工業	95	1,262	12	0	0	0	66	6
	6.厚生・医療	24	96	26	0	0	0	24	0

2003年		A4判	A4変 型判	A5判	A5変 型判	A6判	AB・A5 判	AB判	AB変 型判
	1.総合	299	4,262	294	0	15	12	1,174	64
	2.教育・学芸	24	88	512	12	0	0	60	0
	3.政治・経済・商業	0	468	16	0	0	0	0	0
	4.産業	52	42	0	0	0	0	0	0
	5.工業	101	1,110	12	0	0	0	36	6
	6.厚生・医療	18	96	2	0	0	0	24	0

2004年		A4判	A4変 型判	A5判	A5変 型判	A6判	AB・A5 判	AB判	AB変 型判
	1.総合	312	4,485	288	0	15	12	1,135	64
	2.教育・学芸	12	65	424	12	0	0	48	0
	3.政治・経済・商業	0	444	4	0	0	0	0	0
	4.産業	52	30	0	0	0	0	0	0
	5.工業	118	1,052	12	0	0	0	48	6
	6.厚生・医療	18	112	2	0	0	0	12	0

2005年		A4判	A4変 型判	A5判	A5変 型判	A6判	AB・A5 判	AB判	AB変 型判
	1.総合	343	4,516	277	0	14	12	1,103	64
	2.教育・学芸	13	73	440	12	12	0	48	0
	3.政治・経済・商業	0	445	4	0	0	0	1	0
	4.産業	52	18	0	0	0	0	0	0
	5.工業	131	962	12	0	0	0	48	6
	6.厚生・医療	18	92	2	0	0	0	12	0

B12裁判	B20裁判	B4変型判	B5判	B5変型判	B6判	B6変型判	菊倍判	菊変型判	新書判	null	計
24	12	27	1,357	26	15	5	12	6	0	24	7,585
0	0	0	386	0	0	0	0	0	0	24	1,100
0	0	0	205	1	5	0	0	0	0	0	656
0	0	0	24	0	0	0	0	0	0	0	125
0	0	0	98	12	1	0	0	0	0	0	1,701
0	0	0	45	12	0	0	0	0	0	0	203

B12裁判	B20裁判	B4変型判	B5判	B5変型判	B6判	B6変型判	菊倍判	菊変型判	新書判	null	計
24	12	28	1,344	39	27	5	0	6	12	0	7,651
0	0	0	454	0	0	0	0	0	0	0	1,148
0	0	0	140	12	5	0	0	0	0	0	692
0	0	0	36	0	0	0	1	0	0	0	117
0	0	0	72	0	1	0	0	0	0	0	1,514
0	0	0	57	12	0	0	0	0	0	0	239

B12裁判	B20裁判	B4変型判	B5判	B5変型判	B6判	B6変型判	菊倍判	菊変型判	新書判	null	計
24	12	28	1,380	44	40	3	0	6	12	0	7,668
0	0	0	483	0	0	0	0	0	0	0	1,179
0	0	0	140	0	6	0	0	0	0	0	630
0	0	0	36	0	0	0	1	0	0	0	131
0	0	0	84	0	0	0	0	0	0	0	1,349
0	0	0	70	0	0	0	0	0	0	0	210

B12裁判	B20裁判	B4変型判	B5判	B5変型判	B6判	B6変型判	菊倍判	菊変型判	新書判	null	計
24	12	28	1,225	38	39	3	0	0	12		7,692
0	0	0	443	0	1	0	0	0	0		1,005
0	0	0	139	0	6	0	0	0	0		593
0	0	0	36	0	0	0	1	0	0		119
0	0	0	88	0	0	0	0	0	0		1,324
0	0	0	70	0	0	0	0	0	0		214

B12裁判	B20裁判	B4変型判	B5判	B5変型判	B6判	B6変型判	菊倍判	菊変型判	新書判	null	計
24	12	24	1,308	38	39	3	0	0	12	0	7,789
0	0	0	425	0	1	0	0	0	0	0	1,024
0	0	0	139	0	4	4	0	0	0	0	597
0	0	0	36	0	0	0	1	0	0	0	107
0	0	0	54	0	0	0	0	0	0	0	1,213
0	0	0	82	0	0	0	0	0	0	0	206

表 4.19: 雑誌の文字数調査のまとめ — 発行年・分野・判型別のページ数

⇒ 参照元 3.1.2 節

2001年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	35,396	745,196	58,594	0	4,502	0	230,320	336
	2.教育・学芸	792	8,376	116,980	5,808	0	0	6,516	0
	3.政治・経済・商業	0	60,228	3,624	0	0	0	1,896	0
	4.産業	14,736	5,628	0	0	0	0	0	0
	5.工業	14,228	287,256	3,480	0	0	0	23,316	720
	6.厚生・医療	2,172	15,264	144	0	0	0	4,176	0
2002年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	46,068	763,254	70,508	0	4,502	2,400	198,638	13,336
	2.教育・学芸	624	10,048	108,096	5,808	0	0	6,516	0
	3.政治・経済・商業	0	68,270	3,624	0	0	0	3,948	0
	4.産業	14,352	4,032	0	0	0	0	0	0
	5.工業	13,499	265,146	1,944	0	0	0	16,776	720
	6.厚生・医療	2,172	15,264	5,424	0	0	0	4,176	0
2003年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	46,864	762,612	74,264	0	4,502	2,400	190,066	13,336
	2.教育・学芸	2,160	10,048	118,998	5,808	0	0	6,192	0
	3.政治・経済・商業	0	70,638	4,516	0	0	0	0	0
	4.産業	11,856	6,072	0	0	0	0	0	0
	5.工業	15,803	236,258	1,944	0	0	0	8,832	720
	6.厚生・医療	1,512	15,264	144	0	0	0	4,176	0
2004年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	53,950	816,748	72,776	0	4,502	2,400	186,294	13,336
	2.教育・学芸	624	7,592	100,812	5,808	0	0	4,704	0
	3.政治・経済・商業	0	67,158	1,420	0	0	0	0	0
	4.産業	11,440	4,488	0	0	0	0	0	0
	5.工業	18,896	226,626	1,944	0	0	0	13,392	720
	6.厚生・医療	1,512	17,408	144	0	0	0	2,400	0
2005年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	58,610	822,202	70,896	0	4,322	2,400	183,304	13,336
	2.教育・学芸	642	8,344	102,732	5,808	3,840	0	5,400	0
	3.政治・経済・商業	0	66,598	1,420	0	0	0	136	0
	4.産業	11,440	2,952	0	0	0	0	0	0
	5.工業	21,244	206,540	1,944	0	0	0	13,392	720
	6.厚生・医療	1,512	14,640	144	0	0	0	2,400	0

表 4.20: 雑誌の文字数調査のまとめ — 発行年・分野・判型別の推計総文字数

⇒ 参照元 3.1.2 節

2001年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	36,176,482	761,627,572	60,421,547	0	4,642,417	0	325,598,778	474,996
	2.教育・学芸	606,197	6,410,990	103,499,225	5,138,686	0	0	9,211,539	0
	3.政治・経済・商業	0	61,085,647	3,339,009	0	0	0	2,680,337	0
	4.産業	16,108,806	6,152,304	0	0	0	0	0	0
	5.工業	13,846,974	279,563,284	3,206,333	0	0	0	32,961,363	1,017,850
	6.厚生・医療	2,765,716	19,436,414	132,676	0	0	0	5,903,528	0
2002年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	47,083,799	780,083,751	72,707,145	0	4,642,417	2,922,864	280,810,568	18,852,836
	2.教育・学芸	477,610	7,690,739	95,639,017	5,138,686	0	0	9,211,539	0
	3.政治・経済・商業	0	69,242,165	3,339,009	0	0	0	5,581,209	0
	4.産業	15,689,032	4,407,621	0	0	0	0	0	0
	5.工業	13,137,497	258,045,390	1,791,124	0	0	0	23,715,896	1,017,850
	6.厚生・医療	2,765,716	19,436,414	4,997,457	0	0	0	5,903,528	0
2003年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	47,897,351	779,427,595	76,580,294	0	4,642,417	2,922,864	268,692,503	18,852,836
	2.教育・学芸	1,653,264	7,690,739	105,284,670	5,138,686	0	0	8,753,507	0
	3.政治・経済・商業	0	71,643,885	4,160,862	0	0	0	0	0
	4.産業	12,960,505	6,637,668	0	0	0	0	0	0
	5.工業	15,379,796	229,931,011	1,791,124	0	0	0	12,485,622	1,017,850
	6.厚生・医療	1,925,305	19,436,414	132,676	0	0	0	5,903,528	0
2004年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	55,139,598	834,757,293	75,045,883	0	4,642,417	2,922,864	263,360,102	18,852,836
	2.教育・学芸	477,610	5,810,917	89,194,425	5,138,686	0	0	6,649,951	0
	3.政治・経済・商業	0	68,114,330	1,308,331	0	0	0	0	0
	4.産業	12,505,750	4,906,102	0	0	0	0	0	0
	5.工業	18,389,965	220,556,956	1,791,124	0	0	0	18,932,003	1,017,850
	6.厚生・医療	1,925,305	22,166,477	132,676	0	0	0	3,392,832	0
2005年		A4判	A4変型判	A5判	A5変型判	A6判	AB・A5判	AB判	AB変型判
	1.総合	59,902,351	840,331,554	73,107,246	0	4,456,803	2,922,864	259,133,199	18,852,836
	2.教育・学芸	491,387	6,386,498	90,893,164	5,138,686	3,397,478	0	7,633,872	0
	3.政治・経済・商業	0	67,546,356	1,308,331	0	0	0	192,260	0
	4.産業	12,505,750	3,227,008	0	0	0	0	0	0
	5.工業	20,675,086	201,008,859	1,791,124	0	0	0	18,932,003	1,017,850
	6.厚生・医療	1,925,305	18,641,844	132,676	0	0	0	3,392,832	0

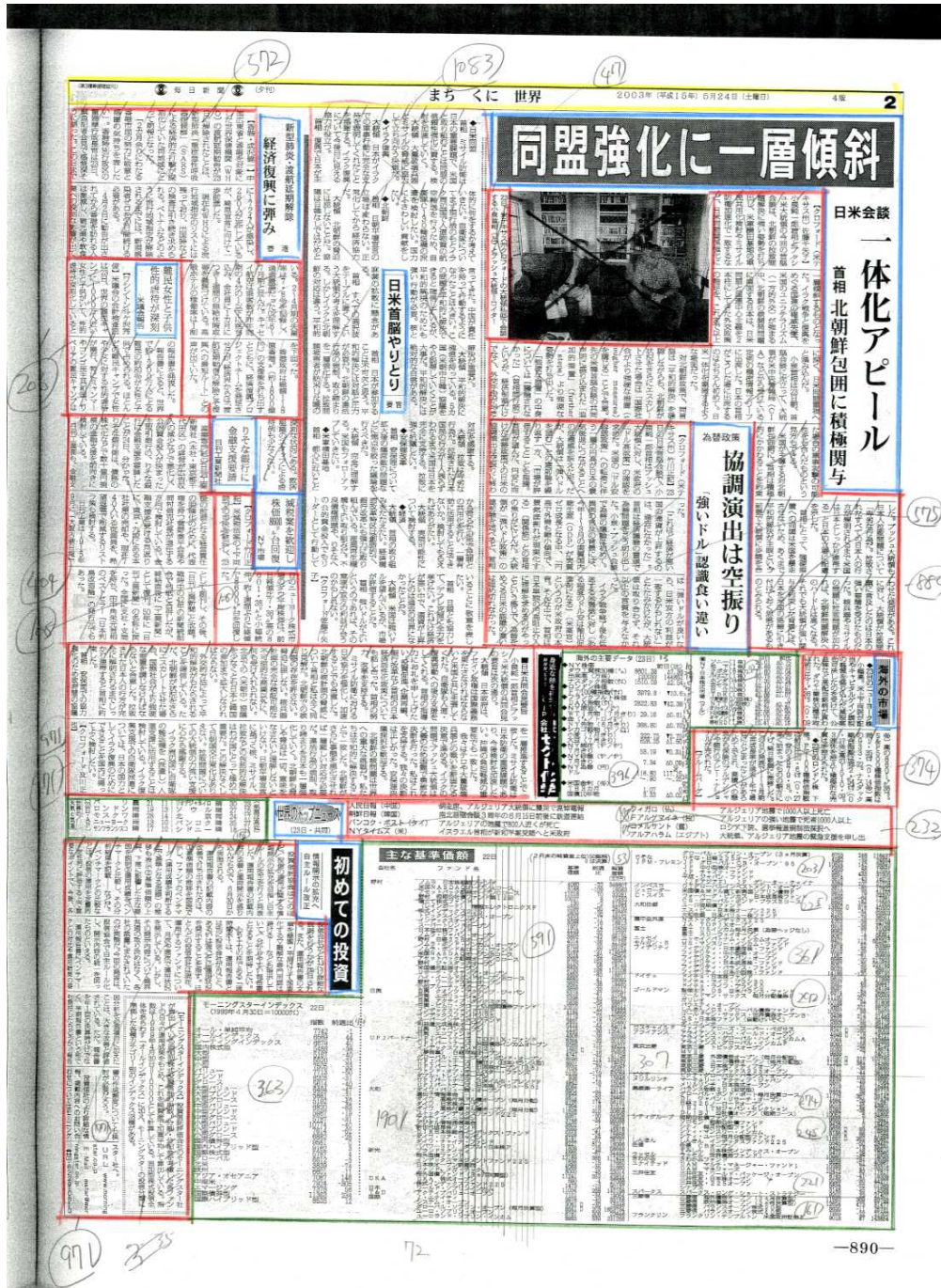


図 4.1: 目視計測で用いた新聞のコピーの例 (その 1)

⇒ 参照元 3.1.3 節

⇒ 参照元 3.1.3 節

表 4.21: 新聞の文字数計測結果 (朝日新聞朝刊)

⇒ 参照元 3.1.3 節

2003.5.12(月)

ページ	面情報	文字数	単純面積	文字数/cm ²
1	一面	5,342	370	14.44
2	総合	5,220	305.25	17.10
3	総合	5,508	305.25	18.04
4	国際	4,974	305.25	16.29
5	経済	3,810	249.75	15.26
6	柱情報なし(教師力)	5,427	305.25	17.78
7	全面広告	43	0	0.00
8	全面広告	43	0	0.00
9	柱情報なし(朝日俳壇・歌壇)	3,961	249.75	15.86
10	オピニオン	5,738	305.25	18.80
11	オピニオン	4,992	305.25	16.35
12	全面広告	46	0	0.00
13	スポーツ(サッカー)	5,133	249.75	20.55
14	全面広告	44	0	0.00
15	スポーツ(相撲)	5,170	305.25	16.94
16	スポーツ(野球)	6,589	305.25	21.59
17	スポーツ(野球)	6,875	397.75	17.28
18	全面広告(見開き)	30	0	0.00
19	全面広告(見開き)	18	0	0.00
20	柱情報なし(くらし)	4,377	305.25	14.34
21	柱情報なし(家庭)	3,527	249.75	14.12
22	全面広告	44	0	0.00
23	全面広告	44	0	0.00
24	全面広告	43	0	0.00
25	全面広告	43	0	0.00
26	柱情報なし(ラジオ・地方局)	8,963	249.75	35.89
27	柱情報なし(BSデジタル・囲碁等)	6,850	305.25	12.36
28	科学・医療	3,108	249.75	12.44
29	全面広告	44	0	0.00
30	全面広告	44	0	0.00
31	東京マリオン	4,988	249.75	19.97
32	全面広告	44	0	0.00
33	東京	5,107	305.25	16.73
34	第2社会	5,188	277.5	18.70
35	社会	4,825	342.25	14.10
36	柱情報なし(地上波テレビ)	11,174	462.5	24.16
計		127,376	6956	-

(cm²)

株式無しのため、株式面のみ2月25日で算出。

2003.2.25付朝刊株式

面	面情報	文字数	単純面積	文字数/cm ²
24	金融情報	22,569	305.25	73.94
25	金融情報	34,261	370	92.60

(cm²)

表 4.22: 新聞の文字数計測結果 (毎日新聞朝刊)

⇒ 参照元 3.1.3 節

2003.3.12(水)

ページ	面情報	文字数	単純面積	文字数/cm ²
1	一面	4,824	351.5	13.72
2	総合 ニュースの焦点	5,603	296	18.93
3	総合 ニュースの焦点	5,094	296	17.21
4	オピニオン ワイド	5,766	296	19.48
5	政治をひらく	6,120	296	20.68
6	全面広告	44	0	0.00
7	人びと 民族 地球	5,468	296	18.47
8	経済がわかる 企業がみえる	6,537	296	22.08
9	経済がわかる 企業がみえる	5,387	296	18.20
10	証券(マザーズ等)	16,764	212.75	78.80
11	証券(東証一部二部)	36,258	453.25	80.00
12	企画特集 ふるさとの主張	3,303	333	9.92
13	主張 提言 討論の広場	4,247	416.25	10.20
14	全面広告	45	0	0.00
15	スポーツ 人間ドラマ(相撲等)	6,316	416.25	15.17
16	スポーツ 人間ドラマ(野球)	7,108	416.25	17.08
17	全面広告	45	0	0.00
18	全面広告	45	0	0.00
19	全面広告	45	0	0.00
20	生活 いきいき 家庭	3,794	333	11.39
21	生活 いきいき 家庭	4,603	416.25	11.06
22	テレビ 見・聞・録 ラジオ(BS&ラジオ)	20,628	453.25	45.51
23	見る 聴く 遊ぶ 東京	5,753	416.25	13.82
24	とうきょう ワイド	6,215	416.25	14.93
25	身近な話題 地域のニュース	4,191	416.25	10.07
26	全面広告	45	0	0.00
27	事件 話題 暮らし	6,119	508.75	11.93
28	社会 事件 ひと 話題	5,418	379.25	14.29
29	社会 事件 ひと 話題	4,994	462.5	10.80
30	テレビ 見・聞・録	11,177	360.75	30.98
計		191,956	8833.75	-

(cm²)

表 4.23: 新聞の文字数計測結果（読売新聞朝刊）

⇒ 参照元 3.1.3 節

2003.4.18(金)

ページ	面情報	文字数	単純面積	文字数/cm ²
1	一面	4,781	351.5	13.60
2	総合	4,983	323	15.43
3	総合	4,881	323	15.11
4	政治	5,810	323	17.99
5	広告のページ	50	0	0.00
6	国際	3,297	266	12.39
7	広告のページ	50	0	0.00
8	広告のページ	50	0	0.00
9	国際	4,416	266	16.60
10	経済	4,544	266	17.08
11	経済	5,089	332.5	15.31
12	経済	4,620	266	17.37
13	広告のページ	53	0	0.00
14	投書	5,594	332.5	16.82
15	解説	5,095	323	15.77
16	スポーツ(野球)	5,989	323	18.54
17	スポーツ(メジャーリーグ)	4,602	332.5	13.84
18	広告のページ	52	0	0.00
19	スポーツ(ワールド)	4,677	332.5	14.07
20	広告のページ	51	0	0.00
21	生活	3,167	256.5	12.35
22	広告のページ	51	0	0.00
23	生活	3,572	332.5	10.74
24	商況(ジャスダック)	20,775	427.5	48.60
25	商況(東証1・2部)	43,019	484.5	88.79
26	広告のページ	55	0	0.00
27	科学	4,143	323	12.83
28	広告のページ	51	0	0.00
29	地域	3,551	266	13.35
30	地域	3,286	256.5	12.81
31	地域	5,034	256.5	19.63
32	広告のページ	50	0	0.00
33	番組(ラジオ・BSデジタル他)	16,292	484.5	33.63
34	社会	5,995	294.5	20.36
35	社会	5,167	361	14.31
36	番組(地上派テレビ・最終面)	11,600	389.5	29.78
計		194,492	8493	—

(cm²)

表 4.24: 新聞の文字数計測結果 (日本経済新聞朝刊)

⇒ 参照元 3.1.3 節

2003.5.27(火)

ページ	面情報	文字数	単純面積	文字数/cm ²
1	一面	5,967	409.5	14.57
2	総合・政治	6,502	341.25	19.05
3	総合	5,476	341.25	16.05
4	全面広告	45	0	0.00
5	経済1	2,076	341.25	6.08
6	全面広告	45	0	0.00
7	経済2	6,009	341.25	17.61
8	全面広告	45	0	0.00
9	特集	4,162	341.25	12.20
10	国際1	5,559	341.25	16.29
11	国際2	5,697	341.25	16.69
12	全面広告	46	0	0.00
13	企業総合	5,349	341.25	15.67
14	企業1	12,113	341.25	35.50
15	企業2	7,075	341.25	20.73
16	企業3	11,107	341.25	32.55
17	企業4 ベンチャー	5,919	341.25	17.35
18	企業財務1	12,353	273	45.25
19	企業財務2	16,605	516.75	32.13
20	全面広告	43	0	0.00
21	全面広告	43	0	0.00
22	マーケット総合1	9,385	341.25	27.50
23	マーケット総合2	19,725	516.75	38.17
24	証券1	34,330	507	67.71
25	証券2	31,621	468	67.57
26	証券3	13,202	341.25	38.69
27	証券4	22,504	341.25	65.95
28	商品	15,322	516.75	29.65
29	経済教室	5,048	273	18.49
30	全面広告	46	0	0.00
31	消費	3,712	364	10.20
32	BS・CSG	6,619	234	28.29
33	首都圏BG(テレビ・ラジオ欄)	17,446	624	27.96
34	全面広告	46	0	0.00
35	首都圏経済・東京	4,426	273	16.21
36	全面広告	46	0	0.00
37	スポーツ	8,543	409.5	20.86
38	社会	4,356	273	15.96
39	社会	6,230	380.25	16.38
40	文化	5,938	341.25	17.40
計		320,781	11498.5	-

(cm²)

表 4.25: 文字数計測を行った新聞の日付と、面積を計測した新聞の日付

⇒ 参照元 3.1.3 節

	曜日	朝日新聞	毎日新聞	読売新聞	日本経済新聞
朝刊	月	◎ 5月12日	5月12日	6月30日	4月14日
	火	2月25日	10月21日	11月11日	◎ 5月27日
	水	4月2日	◎ 3月12日	6月18日	7月9日
	木	4月17日	3月27日	10月2日	11月13日
	金	5月2日	5月16日	◎ 4月18日	3月7日
	土	4月5日	1月4日	1月4日	8月2日
	日	8月31日	11月30日	3月9日	1月5日
夕刊	月	10月20日	6月30日	2月24日	3月31日
	火	10月7日	1月28日	◎ 10月7日	7月1日
	水	2月19日	11月12日	7月23日	10月8日
	木	◎ 10月23日	3月27日	11月13日	12月25日
	金	2月21日	5月16日	5月9日	◎ 1月17日
	土	4月5日	◎ 5月24日	3月1日	6月21日

◎ が付されている日付は、文字数を実測したサンプル。
それ以外は、面種ごとの面積を測定したサンプル。

表 4.26: 全国紙 4 紙の面種のまとめあげ

⇒ 参照元 3.1.3 節

朝日新聞 朝刊

系統	含まれる柱
一面	一面
総合	総合・政治・県議選・憲法を考える・教師力
国際	国際
経済	経済・Weekly経済
特集	朝日俳壇・科学・医療・読書・ノーベル化学賞・読者の写真・NIEスクール・ことばの旅人
オピニオン	オピニオン
スポーツ	スポーツ・サッカー
家庭	家庭・くらし
ラテ欄	ラジオテレビ欄を含む面
地域	東京マリオン・第1～ 東京
社会	第1～3社会・文化総合・明日への環境
証券	株式欄を含む面

読売新聞 朝刊

系統	含まれる柱
一面	一面
総合	総合・政治・解説
国際	国際
経済	経済
特集	News抄録・主張提言・読書・21世紀の選択・科学・衆院選・得票数・議員一覧・歌壇俳壇・人生回廊・小説・追悼抄
オピニオン	投書
スポーツ	スポーツ
家庭	生活・教育・ジュニアプレス・健康
ラテ欄	番組
地域	地域
社会	社会
証券	商況

朝日新聞 夕刊

系統	含まれる柱
一面	一面
総合	世界の鼓動・国内海外
マリオン	マリオン(マリオン・エンターテインメント・Jeans・科学・Let's・芸能)
文芸	芸能・文芸・音楽・コンサート情報・文化・美術展
証券	株式
ラテ欄	ラジオテレビ欄を含む面
特集	おはなしのくに・fashion・惜別・ほがらか
社会	第1～ 社会面
スポーツ	スポーツ

読売新聞 夕刊

系統	含まれる柱
一面	一面
総合	国内海外
いぶにんぐ	いぶにんぐスペシャル
文芸	テレビ情報BOX・文化・芸能・Do! コンポ・モード・本よみうり堂
証券	金融
ラテ欄	テレビラジオ欄を含む面
特集	生命館・写真コンテスト・エコロジー・ズームアップ・安心の設計・人間列島・医療
社会	社会
スポーツ	スポーツ

毎日新聞 朝刊

系統	含まれる柱
一面	一面
総合	総合ニュースの焦点・政治を開く
国際	人々民族地球
経済	経済がわかる企業が見える 本と出会う・短歌俳句・世界学生サミット・企画特集・有事法制・企業アンケート・世論調査・高額ハイカ・アミューズメント・医療・点字文化賞・主張提言討論の広場座談会・メディア・ウィークリー文化
オピニオン	オピニオンワイド
スポーツ	囲碁将棋・スポーツ・テニス
家庭	生活いきいき家庭・学ぶ
ラテ欄	週間TVガイド・テレビ見聞録ラジオ
地域	地域のニュース・見る聴く遊ぶ東京・とうきょうワイド
社会	事件話題暮らし・社会事件ひと話題・環境
証券	証券

日本経済新聞 朝刊

系統	含まれる柱
一面	一面
総合	総合・総合政治・週間予定
経済	経済・経済教室・マネー・ファミリー経済・景気指標・生活経済
特集	特集・スクープ・視点争点・美の美・加齢社
国際	国際・アジア太平洋
企業	企業・企業総合・ベンチャー・企業財務・デジタル・メディア
マーケット	マーケット総合・価格
証券	証券・週間株式・月間株式
商品	商品・消費
ラテ欄	BS・CS、首都圏B
地域	首都圏経済・東京・列島ブラザ、地域総合
スポーツ	スポーツ
社会	社会
文化	文化・読書・芸術・教養・サイエンス・科学・医療・教育

毎日新聞 夕刊

系統	含まれる柱
一面	一面
総合	まちくに世界・選挙
証券	株式欄を含む面
ラテ欄	テレビラジオ欄を含む面
特集	特集ワイド・企画特集・見る写真・学生ウィークリー・テレビのくに・映画囲碁・レジャー・文化・読みごろ
社会	社会
スポーツ	うまウマ馬ランド・スポーツ

日本経済新聞 夕刊

系統	含まれる柱
一面	一面
総合	総合
ビジネス	ビジネス・流通・サービス・ひとビジネス
文化	ひと街・生活・写真特集
マーケット	マーケット総合
証券	証券・商品
ラテ欄	首都圏・タウンビート
スポーツ	スポーツ
社会	社会

表 4.27: 新聞 1 週間分の推計総文字数 (朝日新聞)

⇒ 参照元 3.1.3 節

朝日新聞2003年5月12日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.44	370.00	5,342
総合	17.57	610.50	10,728
国際	16.29	305.25	4,974
経済	15.26	249.75	3,810
特集	15.36	804.75	12,362
オビニオン	17.58	610.50	10,730
スポーツ	19.09	1,258.00	24,015
家庭	14.23	555.00	7,898
ラテ欄	24.14	1,017.50	24,558
地域	18.35	555.00	10,185
社会	16.40	619.75	10,162
証券	83.27	0.00	0
計	22.66	6,956.00	124,764

朝日新聞2003年4月17日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.44	370.00	5,342
総合	17.57	915.75	16,092
国際	16.29	610.50	9,948
経済	15.26	1,008.25	15,381
特集	15.36	240.50	3,694
オビニオン	17.58	610.50	10,730
スポーツ	19.09	832.50	15,892
家庭	14.23	545.75	7,766
ラテ欄	24.14	925.00	22,326
地域	18.35	545.75	10,015
社会	16.40	1,720.50	28,211
証券	83.27	462.50	38,511
計	22.66	8,787.50	183,909

朝日新聞2003年8月31日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.44	370.00	5,342
総合	17.57	610.50	10,728
国際	16.29	305.25	4,974
経済	15.26	610.50	9,313
特集	15.36	1,461.50	22,450
オビニオン	17.58	545.75	9,592
スポーツ	19.09	1,443.00	27,547
家庭	14.23	555.00	7,898
ラテ欄	24.14	1,017.50	24,558
地域	18.35	545.75	10,015
社会	16.40	619.75	10,162
証券	83.27	0.00	0
計	22.66	8,084.50	142,580

朝日新聞2003年5月2日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.44	370.00	5,342
総合	17.57	915.75	16,092
国際	16.29	545.75	8,893
経済	15.26	795.50	12,136
特集	15.36	305.25	4,689
オビニオン	17.58	610.50	10,730
スポーツ	19.09	740.00	14,127
家庭	14.23	545.75	7,766
ラテ欄	24.14	925.00	22,326
地域	18.35	721.50	13,240
社会	16.40	860.25	14,105
証券	83.27	767.75	63,928
計	22.66	8,103.00	193,374

朝日新聞2003年2月25日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.44	370.00	5,342
総合	17.57	1,221.00	21,456
国際	16.29	610.50	9,948
経済	15.26	545.75	8,326
特集	15.36	0.00	0
オビニオン	17.58	610.50	10,730
スポーツ	19.09	1,100.75	21,013
家庭	14.23	545.75	7,766
ラテ欄	24.14	1,008.25	24,335
地域	18.35	555.00	10,185
社会	16.40	1,174.75	19,262
証券	83.27	675.25	56,226
計	22.66	8,417.50	194,589

朝日新聞2003年4月5日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.44	370.00	5,342
総合	17.57	1,285.75	22,594
国際	16.29	545.75	8,893
経済	15.26	804.75	12,277
特集	15.36	795.50	12,220
オビニオン	17.58	610.50	10,730
スポーツ	19.09	1,119.25	21,366
家庭	14.23	305.25	4,344
ラテ欄	24.14	1,008.25	24,335
地域	18.35	804.75	14,768
社会	16.40	925.00	15,167
証券	83.27	832.50	69,320
計	22.66	9,407.25	221,355

朝日新聞2003年4月2日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.44	370.00	5,342
総合	17.57	915.75	16,092
国際	16.29	555.00	9,044
経済	15.26	795.50	12,136
特集	15.36	619.75	9,520
オビニオン	17.58	610.50	10,730
スポーツ	19.09	1,230.25	23,485
家庭	14.23	555.00	7,898
ラテ欄	24.14	1,008.25	24,335
地域	18.35	555.00	10,185
社会	16.40	1,119.25	18,352
証券	83.27	767.75	63,928
計	22.66	9,102.00	211,047

表 4.28: 新聞 1 週間分の推計総文字数 (毎日新聞)

⇒ 参照元 3.1.3 節

毎日新聞2003年3月12日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.72	351.50	4,824
総合	18.94	888.00	16,817
国際	18.47	296.00	5,468
経済	20.14	592.00	11,924
特集	10.06	749.25	7,538
オビニオン	19.48	296.00	5,766
スポーツ	16.12	832.50	13,424
家庭	11.23	749.25	8,411
ラテ欄	38.25	814.00	31,133
地域	12.94	1,248.75	16,159
社会	12.34	1,350.50	16,664
証券	79.40	666.00	52,878
計	22.59	8,833.75	191,006

毎日新聞2003年3月27日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.72	351.50	4,824
総合	18.94	888.00	16,817
国際	18.47	536.50	9,911
経済	20.14	592.00	11,924
特集	10.06	999.00	10,051
オビニオン	19.48	296.00	5,766
スポーツ	16.12	1,017.50	16,407
家庭	11.23	656.75	7,373
ラテ欄	38.25	814.00	31,133
地域	12.94	832.50	10,773
社会	12.34	592.00	7,305
証券	79.40	666.00	52,878
計	22.59	8,241.75	185,161

毎日新聞2003年11月30日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.72	351.50	4,824
総合	18.94	592.00	11,211
国際	18.47	296.00	5,468
経済	20.14	296.00	5,962
特集	10.06	1,905.50	19,171
オビニオン	19.48	296.00	5,766
スポーツ	16.12	1,674.25	26,997
家庭	11.23	296.00	3,323
ラテ欄	38.25	1,720.50	65,804
地域	12.94	296.00	3,830
社会	12.34	592.00	7,305
証券	79.40	0.00	0
計	22.59	8,315.75	159,661

毎日新聞2003年5月12日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.72	351.50	4,824
総合	18.94	592.00	11,211
国際	18.47	536.50	9,911
経済	20.14	240.50	4,844
特集	10.06	1,128.50	11,354
オビニオン	19.48	296.00	5,766
スポーツ	16.12	1,498.50	24,163
家庭	11.23	296.00	3,323
ラテ欄	38.25	814.00	31,133
地域	12.94	592.00	7,661
社会	12.34	888.00	10,957
証券	79.40	0.00	0
計	22.59	7,233.50	125,147

毎日新聞2003年5月16日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.72	351.50	4,824
総合	18.94	888.00	16,817
国際	18.47	296.00	5,468
経済	20.14	1,110.00	22,358
特集	10.06	434.75	4,374
オビニオン	19.48	296.00	5,766
スポーツ	16.12	693.75	11,187
家庭	11.23	656.75	7,373
ラテ欄	38.25	814.00	31,133
地域	12.94	536.50	6,942
社会	12.34	943.50	11,642
証券	79.40	453.25	35,986
計	22.59	7,474.00	163,869

毎日新聞2003年10月21日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.72	351.50	4,824
総合	18.94	888.00	16,817
国際	18.47	536.50	9,911
経済	20.14	804.75	16,209
特集	10.06	1,008.25	10,144
オビニオン	19.48	296.00	5,766
スポーツ	16.12	749.25	12,082
家庭	11.23	536.50	6,023
ラテ欄	38.25	814.00	31,133
地域	12.94	888.00	11,491
社会	12.34	629.00	7,761
証券	79.40	453.25	35,986
計	22.59	7,955.00	168,147

毎日新聞2003年1月4日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.72	351.50	4,824
総合	18.94	592.00	11,211
国際	18.47	536.50	9,911
経済	20.14	0.00	0
特集	10.06	1,480.00	14,890
オビニオン	19.48	0.00	0
スポーツ	16.12	1,655.75	26,699
家庭	11.23	296.00	3,323
ラテ欄	38.25	814.00	31,133
地域	12.94	296.00	3,830
社会	12.34	592.00	7,305
証券	79.40	0.00	0
計	22.59	6,613.75	113,126

表 4.29: 新聞1週間分の推計総文字数(読売新聞)

⇒ 参照元 3.1.3 節

読売新聞2003年4月18日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.60	351.50	4,781
総合	16.08	1,292.00	20,769
国際	14.50	532.00	7,713
経済	16.59	864.50	14,338
特集	12.83	323.00	4,143
オピニオン	16.82	332.50	5,594
スポーツ	15.48	988.00	15,297
家庭	11.54	589.00	6,800
ラテ欄	31.70	874.00	27,709
地域	15.26	779.00	11,889
社会	17.33	655.50	11,363
証券	68.69	912.00	62,648
計	20.87	8,493.00	193,045

読売新聞2003年10月2日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.60	389.50	5,298
総合	16.08	1,292.00	20,769
国際	14.50	589.00	8,539
経済	16.59	855.00	14,181
特集	12.83	0.00	0
オピニオン	16.82	323.00	5,434
スポーツ	15.48	978.50	15,150
家庭	11.54	532.00	6,142
ラテ欄	31.70	874.00	27,709
地域	15.26	798.00	12,179
社会	17.33	655.50	11,363
証券	68.69	912.00	62,648
計	20.87	8,198.50	189,413

読売新聞2003年6月30日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.60	389.50	5,298
総合	16.08	969.00	15,577
国際	14.50	323.00	4,683
経済	16.59	646.00	10,714
特集	12.83	1,121.00	14,379
オピニオン	16.82	323.00	5,434
スポーツ	15.48	1,615.00	25,005
家庭	11.54	1,387.00	16,013
ラテ欄	31.70	874.00	27,709
地域	15.26	323.00	4,930
社会	17.33	617.50	10,704
証券	68.69	0.00	0
計	20.87	8,588.00	140,445

読売新聞2003年1月4日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.60	389.50	5,298
総合	16.08	1,292.00	20,769
国際	14.50	285.00	4,132
経済	16.59	323.00	5,357
特集	12.83	323.00	4,143
オピニオン	16.82	0.00	0
スポーツ	15.48	1,235.00	19,121
家庭	11.54	589.00	6,800
ラテ欄	31.70	874.00	27,709
地域	15.26	323.00	4,930
社会	17.33	655.50	11,363
証券	68.69	0.00	0
計	20.87	6,289.00	109,622

読売新聞2003年11月11日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.60	389.50	5,298
総合	16.08	1,292.00	20,769
国際	14.50	589.00	8,539
経済	16.59	589.00	9,769
特集	12.83	3,496.00	44,842
オピニオン	16.82	323.00	5,434
スポーツ	15.48	912.00	14,120
家庭	11.54	532.00	6,142
ラテ欄	31.70	874.00	27,709
地域	15.26	646.00	9,859
社会	17.33	655.50	11,363
証券	68.69	912.00	62,648
計	20.87	11,210.00	226,494

読売新聞2003年3月9日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.60	389.50	5,298
総合	16.08	1,292.00	20,769
国際	14.50	589.00	8,539
経済	16.59	323.00	5,357
特集	12.83	1,558.00	19,984
オピニオン	16.82	323.00	5,434
スポーツ	15.48	1,235.00	19,121
家庭	11.54	532.00	6,142
ラテ欄	31.70	1,748.00	55,419
地域	15.26	589.00	8,989
社会	17.33	655.50	11,363
証券	68.69	0.00	0
計	20.87	9,234.00	166,416

読売新聞2003年6月18日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	13.60	389.50	5,298
総合	16.08	1,292.00	20,769
国際	14.50	589.00	8,539
経済	16.59	855.00	14,181
特集	12.83	589.00	7,555
オピニオン	16.82	323.00	5,434
スポーツ	15.48	1,035.50	16,032
家庭	11.54	855.00	9,871
ラテ欄	31.70	874.00	27,709
地域	15.26	855.00	13,049
社会	17.33	655.50	11,363
証券	68.69	912.00	62,648
計	20.87	9,224.50	202,449

表 4.30: 新聞 1 週間分の推計総文字数（日本経済新聞）

⇒ 参照元 3.1.3 節

日本経済新聞2003年1月5日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.57	409.50	5,967
総合	17.55	682.50	11,978
経済	14.06	1,706.25	23,992
特集	12.20	1,706.25	20,810
国際	16.49	341.25	5,628
企業	28.45	341.25	9,710
マーケット	32.84	0.00	0
証券	59.98	0.00	0
商品	19.92	0.00	0
ラテ欄	28.12	926.25	26,048
地域	16.21	0.00	0
スポーツ	20.86	341.25	7,119
社会	16.17	653.25	10,563
文化	17.40	2,388.75	41,566
計	22.49	9,496.50	163,381

日本経済新聞2003年5月27日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.57	409.50	5,967
総合	17.55	682.50	11,978
経済	14.06	955.50	13,435
特集	12.20	341.25	4,162
国際	16.49	682.50	11,256
企業	28.45	2,496.00	71,022
マーケット	32.84	858.00	28,174
証券	59.98	1,657.50	99,413
商品	19.92	880.75	17,548
ラテ欄	28.12	858.00	24,129
地域	16.21	273.00	4,426
スポーツ	20.86	409.50	8,543
社会	16.17	653.25	10,563
文化	17.40	341.25	5,938
計	22.49	11,498.50	316,554

日本経済新聞2003年7月9日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.57	409.50	5,967
総合	17.55	682.50	11,978
経済	14.06	955.50	13,435
特集	12.20	0.00	0
国際	16.49	682.50	11,256
企業	28.45	2,086.50	59,370
マーケット	32.84	858.00	28,174
証券	59.98	1,716.00	102,922
商品	19.92	789.75	15,735
ラテ欄	28.12	653.25	18,371
地域	16.21	273.00	4,426
スポーツ	20.86	516.75	10,780
社会	16.17	653.25	10,563
文化	17.40	341.25	5,938
計	22.49	10,617.75	298,915

日本経済新聞2003年4月14日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.57	409.50	5,967
総合	17.55	955.50	16,769
経済	14.06	955.50	13,435
特集	12.20	0.00	0
国際	16.49	955.50	15,758
企業	28.45	1,296.75	36,898
マーケット	32.84	273.00	8,964
証券	59.98	0.00	0
商品	19.92	0.00	0
ラテ欄	28.12	653.25	18,371
地域	16.21	273.00	4,426
スポーツ	20.86	789.75	16,476
社会	16.17	721.50	11,667
文化	17.40	955.50	16,626
計	22.49	8,238.75	165,358

日本経済新聞2003年8月2日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.57	409.50	5,967
総合	17.55	682.50	11,978
経済	14.06	955.50	13,435
特集	12.20	0.00	0
国際	16.49	682.50	11,256
企業	28.45	1,745.25	49,660
マーケット	32.84	858.00	28,174
証券	59.98	4,299.75	257,890
商品	19.92	516.75	10,296
ラテ欄	28.12	653.25	18,371
地域	16.21	273.00	4,426
スポーツ	20.86	516.75	10,780
社会	16.17	653.25	10,563
文化	17.40	614.25	10,688
計	22.49	12,860.25	443,484

日本経済新聞2003年3月7日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.57	409.50	5,967
総合	17.55	682.50	11,978
経済	14.06	614.25	8,637
特集	12.20	175.50	2,140
国際	16.49	682.50	11,256
企業	28.45	2,047.50	58,260
マーケット	32.84	858.00	28,174
証券	59.98	1,677.00	100,583
商品	19.92	789.75	15,735
ラテ欄	28.12	653.25	18,371
地域	16.21	273.00	4,426
スポーツ	20.86	409.50	8,543
社会	16.17	653.25	10,563
文化	17.40	341.25	5,938
計	22.49	10,266.75	290,571

日本経済新聞2003年11月13日 朝刊

面種	文字数/cm ²	面積合計	文字数
一面	14.57	409.50	5,967
総合	17.55	682.50	11,978
経済	14.06	955.50	13,435
特集	12.20	0.00	0
国際	16.49	682.50	11,256
企業	28.45	1,774.50	50,492
マーケット	32.84	858.00	28,174
証券	59.98	1,686.75	101,168
商品	19.92	789.75	15,735
ラテ欄	28.12	653.25	18,371
地域	16.21	273.00	4,426
スポーツ	20.86	409.50	8,543
社会	16.17	653.25	10,563
文化	17.40	341.25	5,938
計	22.49	10,169.25	286,046

表 4.31: 全国紙 4 紙 1 週間分の文字数・ページ数と、キャラクタ密度の算出

⇒ 参照元 3.1.3 節

朝日新聞	朝刊		夕刊	
	ページ数	文字数	ページ数	文字数
月	36	124,764	20	93,641
火	36	194,589	16	113,091
水	40	211,046	16	80,221
木	36	183,908	24	107,259
金	32	193,373	20	95,063
土	36	221,355	16	102,251
日	40	142,579	—	—
計	256	1,271,614	112	591,526
1日平均	37	181,659	19	98,588
キャラクタ密度	4,967		5,281	

毎日新聞	朝刊		夕刊	
	ページ数	文字数	ページ数	文字数
月	28	125,146	10	79,908
火	32	168,146	10	79,656
水	30	191,005	10	79,355
木	32	185,160	14	100,040
金	30	163,869	12	98,668
土	28	113,126	10	65,269
日	32	159,661	—	—
計	212	1,106,113	66	502,896
1日平均	30	158,016	11	83,816
キャラクタ密度	5,218		7,620	

読売新聞	朝刊		夕刊	
	ページ数	文字数	ページ数	文字数
月	36	140,445	24	100,746
火	40	226,493	16	99,601
水	40	202,449	20	107,618
木	36	189,412	20	98,424
金	36	193,045	20	100,851
土	32	109,622	16	111,350
日	40	166,415	—	—
計	260	1,227,881	116	618,590
1日平均	37	175,412	19	103,098
キャラクタ密度	4,723		5,333	

日本経済新聞	朝刊		夕刊	
	ページ数	文字数	ページ数	文字数
月	40	165,357	24	132,192
火	40	316,554	16	151,266
水	40	298,915	18	145,876
木	44	286,045	18	145,876
金	44	290,570	16	150,076
土	36	443,483	12	80,437
日	32	163,381	—	—
計	276	1,964,305	104	805,723
1日平均	39	280,615	17	134,287
キャラクタ密度	7,117		7,747	

表 4.32: 新聞の文字数調査のまとめ — 新聞タイトル（朝夕刊）別の推計総文字数

⇒ 参照元 3.1.3 節

紙種	新聞タイトル		総ページ数	キャラクタ密度	総文字数
全国	朝日新聞	朝刊	64,914	4967.24	322,444,837
		夕刊	26,413	5281.48	139,501,580
	毎日新聞	朝刊	53,757	5217.51	280,478,468
		夕刊	15,565	7619.64	118,599,697
	読売新聞	朝刊	65,929	4722.62	311,355,489
		夕刊	27,357	5332.67	145,883,987
	日本経済新聞	朝刊	69,986	7117.05	498,091,726
		夕刊	24,527	7747.34	190,016,297
	産経新聞	朝刊	56,800	4969.13	282,246,311
夕刊		21,225	6077.93	129,004,071	
ブロック	北海道新聞	朝刊	63,900	4969.13	317,527,100
		夕刊	16,980	6077.93	103,203,257
	中日新聞	朝刊	63,900	4969.13	317,527,100
		夕刊	22,640	6077.93	137,604,342
	西日本新聞	朝刊	63,900	4969.13	317,527,100
		夕刊	16,980	6077.93	103,203,257
地方	河北新報	朝刊	60,350	4969.13	299,886,705
		夕刊	11,320	6077.93	68,802,171
	新潟日報	朝刊	56,800	4969.13	282,246,311
		夕刊	11,320	6077.93	68,802,171
	京都新聞	朝刊	51,954	4969.13	258,167,172
		夕刊	16,386	6077.93	99,591,143
	神戸新聞	朝刊	53,836	4969.13	267,516,581
		夕刊	18,041	6077.93	109,653,460
	中国新聞	朝刊	56,800	4969.13	282,246,311
		夕刊	11,320	6077.93	68,802,171
	高知新聞	朝刊	46,150	4969.13	229,325,128
		夕刊	11,320	6077.93	68,802,171
	愛媛新聞	朝刊	49,700	4969.13	246,965,522
		夕刊	11,320	6077.93	68,802,171
琉球新報	朝刊	56,800	4969.13	282,246,311	
	夕刊	11,320	6077.93	68,802,171	
合計			1,198,189		6,416,070,114

研究開発部門 言語資源グループ（サンプリングサブグループ）

山崎 誠	（国立国語研究所副グループ長）
柏野 和佳子	（国立国語研究所研究員）
丸山 岳彦*	（国立国語研究所研究員）
佐野 大樹	（国立国語研究所特別奨励研究員）
秋元 祐哉*	（国立国語研究所研究補佐員）
稲益 佐知子	（国立国語研究所研究補佐員）
吉田谷 幸宏	（国立国語研究所研究補佐員）

（* は執筆者）

国立国語研究所内部報告書（LR-CCG-06-02）

『現代日本語書き言葉均衡コーパス』におけるサンプル構成比の算出法

—現代日本語書き言葉の文字数調査—

平成 19 年 3 月 31 日

執筆者 丸山 岳彦 秋元 祐哉

発行者 独立行政法人国立国語研究所

連絡先 〒 190-8561 東京都立川市緑町 10-2

電話 042 (540) 4300（代表）



国立国語研究所

